ENCICLOPEDIA DE

LAS CIENCIAS

ENCICLOPEDIA DE
LAS CIENCIAS

VOLUMEN 2

La exploración de la Luna

Computadoras y matemáticas

Ciencia de la Tierra I

Editorial Cumbre, S.A.

UNA EMPRESA GROLIER

Sexta edición
© 1983 Editorial Cumbre, S.A.
Crepúsculo 46
Col. Insurgentes Cuicuilco
México, D.F.

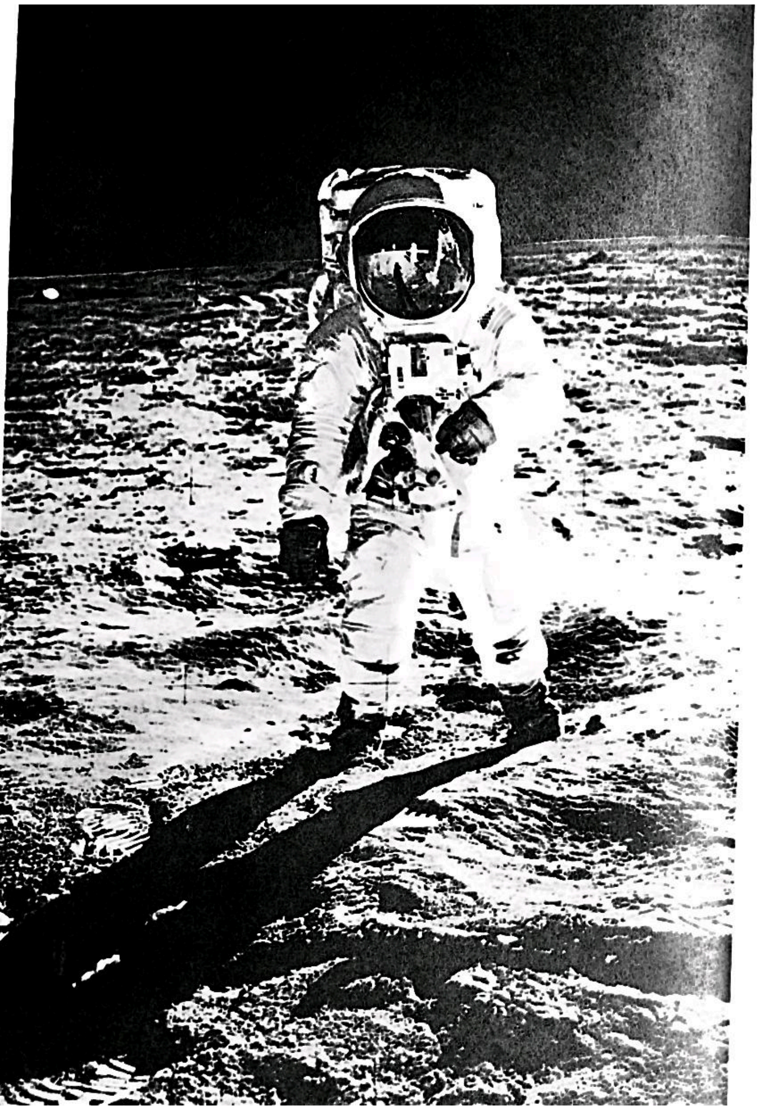
Reservados todos los derechos bajo la Convención Universal, los Convenios Panamericanos y el Convenio de la Unión Interamericana de Derechos de Autor.

ISBN 968-33-0052-9 edición completa
ISBN 968-33-0054-5 volumen 2
Impreso en México
Printed in Mexico

Volumen 2

Contenido

La exploración de la Luna	1 - 35
Computadoras y matemáticas	36 - 191
Ciencia de la Tierra (1a. parte)	192 - 296



LA EXPLORACIÓN DE LA LUNA

La Luna siempre ha interesado al hombre. Desde los tiempos primigenios éste ha especulado sobre su naturaleza y origen. Durante los siglos en que la Luna fue tema de mitos y supersticiones, no podía hacerse ningún intento serio para entenderla y alcanzarla. Primero se debía reconocer que era un mundo espacial diferente, aunque semejante a la Tierra en algunos aspectos. Segundo, era menester darse cuenta de las dificultades implicadas en un posible viaje al satélite.

Los antiguos griegos fueron los primeros en especular libremente sobre el Universo y en darse cuenta de que la Tierra era un globo y que el Sol, la Luna y los planetas eran, quizá, mundos como la Tierra. Ya antes los pueblos del antiguo Oriente habían acumulado una enorme cantidad de información sobre los movimientos aparentes de estas luces celestes, pero no las reconocieron como cuerpos siderales; se necesitó un largo período de especulación, medición y estudio antes de alcanzar el estado actual de conocimiento y exploración lunares.

La historia del estudio y la exploración de la Luna se divide en tres períodos principales: (1) La era pretelescópica, hasta 1609; (2) La era telescópica, desde 1609 hasta 1959, y (3) La era de la exploración espacial, iniciada en 1959, año en que la primera sonda lunar alcanzó el satélite. En 1969, los primeros hombres pusieron el pie en él, acontecimiento que marcó otro hito en su exploración.

EXPLORACIÓN PRETELESCÓPICA

Tal como hemos dicho, los antiguos griegos sugirieron que podría haber otros mundos en el espacio. Midieron la circunferencia de la Tierra con sorprendente precisión. Empero, sus estimaciones del tamaño de la Luna y de su distancia de la Tierra no fueron muy precisas. Al carecer de telescopios e instrumentos similares sólo podían conjeturar sobre su naturaleza. En general, los griegos creían que la Tierra era el centro del Universo, y que el Sol, la Luna, los planetas y las estrellas giraban en torno de nosotros. Sin embargo, el astrónomo griego Aristarco de Samos propuso, en el siglo III a.C., que el Sol era el centro de las revoluciones planetarias, y no la Tierra. Esta idea fue, en general, desatendida hasta el siglo XVI, cuando la exhumó exitosamente el astrónomo polaco Nicolás Copérnico.

Ante la carencia de cualquier conocimiento exacto sobre la atmósfera y el espacio exterior, algunos griegos consideraron posible alcanzar la Luna haciéndose a la vela, o volando. El primer relato conocido de un viaje a la Luna fue escrito por un griego llamado Luciano de Samosata, en el siglo II; narra la aventura de una nave en el Atlántico, impulsada hacia el satélite por una fuerte tormenta. Allí, los

Esta antena de televisión, unida al Lunar Rover, vehículo de los astronautas, permitió a los científicos explorar la Luna desde la Tierra juntamente con aquéllos.

NASA

Vista panorámica de la superficie lunar. Atravesándola en diagonal están las huellas dejadas en el polvo por un dispositivo trasportador, que se alejaba del módulo (al fondo).

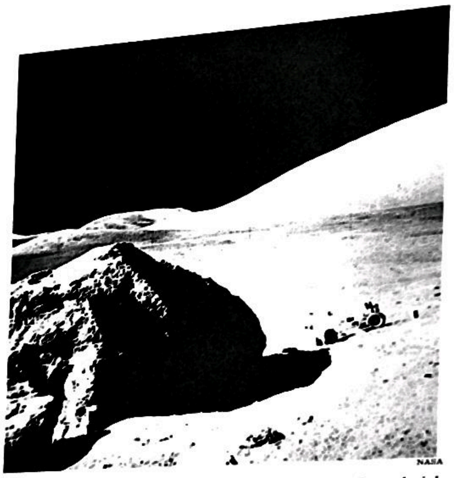
pasajeros se encuentran con seres fantásticos que están preparando una guerra contra el Sol. Luciano escribió otro relato, sobre un hombre llamado Ícaro, que vuela a la Luna mediante alas de ave adheridas a su cuerpo. Estas narraciones, básicamente satíricas, fijan el patrón para muchas otras de los siglos siguientes. Se creyó que la Luna y otros mundos eran lares de seres parecidos al hombre, de monstruos y otras criaturas horripilantes o fuera de lo común.

En la Edad Media europea (500-1500), la Iglesia católica cristalizó sus enseñanzas acerca del Universo sobre el sistema tolemaico, según el cual la Tierra era el centro del Cosmos. La especulación sobre la vida en otros mundos fue considerada peligrosa, y se la desalentó. Sin embargo, tiempo después, las investigaciones comenzaron a revelar los errores de estas ideas; en 1543, Copérnico repitió el punto de vista de Aristarco, para quien la Tierra y otros planetas giraban en torno del Sol. En 1609, el científico italiano Galileo dirigió el telescopio recientemente inventado hacia los cielos y lo que allí vio asestó un golpe mortal al sistema tolemaico, iniciando una revolución que estableció los fundamentos de la astronomía moderna.

LA ERA TELESCÓPICA

El pequeño y tosco telescopio empleado por Galileo mostró la faz de la Luna como nunca se había visto antes. Incluso su modesta amplificación de unas treinta veces reveló montañas brillantes, cráteres y oscuras áreas llanas. Galileo trazó una carta rudimentaria de las principales características de la Luna vista a través de su instrumento, y quedó tan impresionado por su aspecto aparentemente terrestre que erróneamente llamó a las regiones oscuras *maria*, palabra latina para "mares".

En el siglo XVII siguieron otros descubrimientos astronómicos. Johannes Kepler, astrónomo alemán, formuló las leyes del movimiento planetario; una establece que los planetas describen órbitas en torno del Sol, pero éstas no son circulares, como había creído Copérnico, sino trayectorias ovales llamadas elipses. El físico inglés Isaac Newton descubrió las leyes de la gravitación, que ayudaron a explicar el movimiento planetario. Entonces se produjo una proliferación de relatos sobre viajes a la Luna, algunos de los cuales incorporaron los nuevos descubrimientos; empero, la mayoría era fantasía, donde los medios para viajar a la Luna eran totalmente absurdos, excepto en un relato en el cual se empleaba un vehículo impulsado por cohetes.



El Lunar Rover empleado por los astronautas de la Apolo 17 queda empequeñecido por un peñasco gigantesco en Taurus-Littrow, región montañosa de la Luna.

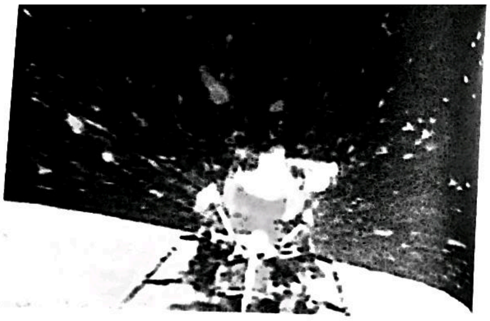
A medida que avanzaba el siglo XVII, los astrónomos se daban cuenta cada vez más de los inmensos abismos que separan Tierra, Luna, Sol y planetas. La Luna se suponía un globo sin aire y pequeño, incapaz de albergar vida terrestre.

En los siglos XVIII y XIX se produjeron formidables progresos en astronomía; los telescopios aumentaron enormemente su poder y claridad, y nuevos instrumentos –como el espectroscopio y la cámara fotográfica– expandieron enormemente los conocimientos. El advenimiento de la fotografía en el siglo XIX facilitó y aceleró la tarea de trazar un mapa lunar; la primera fotografía fue tomada en 1840 por un científico norteamericano, John Draper. En 1897, el Observatorio de París completó un atlas fotográfico del satélite.

Para el siglo XIX, la astronomía y la geología habían avanzado suficientemente, de modo que podían ofrecerse explicaciones serias de los caracteres lunares, y tratar sobre los orígenes de éstos.

El siglo XX ha sido testigo de revoluciones aún mayores en astronomía. Mediante el uso del telescopio se han podido observar detalles de la superficie lunar que el ojo humano nunca había visto antes, y estos descubrimientos decidieron a pequeños grupos de personas a considerar seriamente medios para viajar a la Luna. Mientras tanto, los escritores de ficción no habían estado inactivos; varios autores, entre ellos Julio Verne y H. G. Wells, mantuvieron vivo, de tiempo en tiempo, el interés popular, con sus relatos de aventuras interplanetarias y lunares.

Si bien desde el siglo XIII los cohetes se usaban como fuegos de artificio y como armas, pocos individuos alguna vez los consideraron como un medio seguro, o confiable, para viajar. Sólo al llegar el siglo XIX se hicieron propuestas serias e, incluso, experi-



El módulo lunar del Apolo 16 fue llamado *Orión*. A medida que despega de la Luna, la llama del motor de su etapa de ascenso crea un efecto caleidoscópico.

NASA

mentos, utilizando vehículos propulsados por cohetes. Hacia fines del siglo XIX, el científico ruso Constantin Tsiolkovsky y el científico alemán Hermann Ganswindt escribieron y disertaron sobre los principios del vuelo espacial autopropulsado, pero, en ese momento, pocos científicos e ingenieros tomaron muy en serio las ideas de estos dos hombres.

El siglo XX vivió la realización del sueño de la nave cohete; en las décadas del 20 y del 30 se colocaron los cimientos de la tecnología que por fin llevó al hombre y sus instrumentos a la Luna, y más allá aún. Después de la Segunda Guerra Mundial –que vio el desarrollo de la tecnología cohetera–, y hacia fines de la década del 40 y durante la del 50, la Unión Soviética y los Estados Unidos lanzaron cohetes para explorar la atmósfera superior y el espacio próximo. Entonces, decidieron lanzar satélites artificiales para circundar la Tierra y, después, para que giraran en órbita lunar, o alunizaran.

EXPLORACIÓN ESPACIAL DE LA LUNA

En 1957, la Unión Soviética lanzó el Sputnik, primer satélite artificial que giró en torno de la Tierra. Inmediatamente la siguieron los Estados Unidos con el Explorer 1, a comienzos de 1958. Luego, la Unión Soviética envió a la Luna sondas no tripuladas, en su serie de Luniks. Después, en 1959, el Lunik 2 se estrelló en la Luna. En el mismo año, el Lunik 3 la circundó y envió a la Tierra imágenes televisadas de su cara oculta, que nunca había sido vista antes por el hombre. En 1966 el Lunik 9 efectuó un alunizaje suave sobre la superficie lunar y trasmitió imágenes a la Tierra. Mientras tanto, los Estados Unidos planearon enviar sondas no tripuladas a la Luna. Su serie de naves lunares Ranger, en 1964 y 1965, fue concebida para estrellarse en el satélite, pero antes, y a una señal de la base terrestre, el Ranger tomaba miles de fotos de la superficie lunar mientras se aproximaba. Las fotografías eran trasmitidas electrónicamente a la Tierra. Tres naves Ranger, 7-9, mostraron mil veces más detalles de la superficie lunar que los que se podían ver con los mejores telescopios terrestres; incluso se hicieron visibles pequeños cráteres de sólo uno o dos metros de diámetro.

En 1966 y 1967 los Estados Unidos lanzaron satélites artificiales no tripulados para que girasen en órbita lunar. Conocidos como Lunar Orbiter, cinco de ellos pasaron muy cerca de la superficie de la Luna y tomaron miles de fotografías, que fueron retrasmitidas a la Tierra. Asimismo, estos vehículos tomaron fotos de la Tierra tal como se ve desde la Luna. Hubo otra serie de lanzamientos de naves no tripuladas concebidas para descender con suavidad sobre la Luna, llamadas Surveyor 1-7; cinco de ellas se posaron suavemente en diversos puntos del satélite para probar su superficie. Determinaron que se podía descender en él sin riesgo. También, llevaron a cabo diversas pruebas en la superficie lunar y excavaron algo de ella para su estudio. El Surveyor 5 hizo un ensayo químico de rocas lunares, irradiándolas con partículas atómicas y registrando cómo reaccionaban. Así, los científicos en la Tierra dedujeron algo de la composición química del material lunar; concluyeron que era, como el basalto terrestre, una roca volcánica.

Los Surveyor mostraron que la superficie de la Luna estaba cubierta con roca pulverizada y trozos más grandes de piedra. Asimismo enviaron muchas fotos de la Luna a la Tierra, y también imágenes de la Tierra como se la ve desde la Luna.

En las décadas de 1960 y 1970, la Unión Soviética continuó su programa de exploración lunar no tripulada. Las sondas circundaban la Luna, o aterrizaban en su superficie. Las sondas Luna tomaron muestras del suelo lunar y luego las enviaron, por medio de cohetes, a la Tierra para su estudio. Se hicieron descender dos vehículos con ruedas –Lunojod 1 y 2– sobre la superficie de la Luna, y se los guió por control remoto desde la Tierra.

Ya en 1961, en el Proyecto Apolo, los Estados Unidos habían decidido hacer descender hombres en el satélite. Desde fines de 1968, una serie de misiones Apolo envió astronautas de E.U. para ponerlos en órbita en torno de la Luna y, luego, los hizo descender en su superficie. Los astronautas trajeron muchos especímenes de roca que acrecentaron inmensamente nuestro caudal de conocimientos sobre el satélite; también tomaron miles de fotografías de la Luna y de la Tierra, en la misma Luna y desde la nave espacial Apolo, y realizaron experimentos científicos sobre la superficie lunar, cuyos resultados fueron enviados por televisión a la Tierra.

VIAJE EN LA SUPERFICIE DE LA LUNA

La Luna presenta muchos obstáculos para viajar sobre su superficie o por encima de ella. Las modalidades convencionales de trasporte terrestre –aeroplano, automóvil, tren y barco– serían muy difíciles, o imposibles, allí, pues la Luna no tiene aire. Los motores de combustión, como los diesel, de gasolina, de vapor y de chorro, no pueden operar, a menos que se suministre aire u oxígeno y combustible. La ausencia de atmósfera en la Luna también descarta los aviones. Sólo motores a reacción o cohetes podrían hacer posible el vuelo a través de los cielos lunares. No hay masas de agua para barcos.

La locomoción en la superficie lunar con un vehículo que se desplace sobre el suelo implica muchos problemas. Los motores que no requieren aire serían impulsados con electricidad. La baja gravedad lunar –un sexto de la que posee la superficie de la Tierra– dificultaría el control del vehículo; tendrían que diseñarse sistemas especiales de suspensión, dirección y ejes.

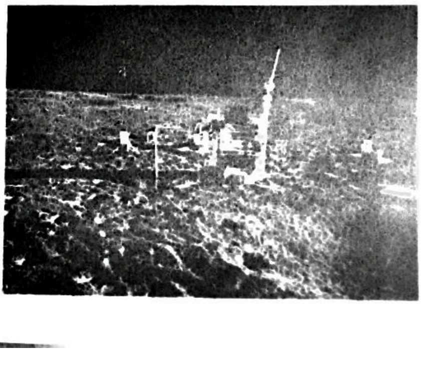
¿Qué clase de tren de rodamiento sería apto? La tracción, o agarre, sobre el polvo lunar, blando y flojo, sería difícil de lograr. Los neumáticos con aire no resultarían; simplemente, se inflarían y explotarían en el vacío de la Luna. Una solución serían las cadenas o cremalleras, como en un tanque; pero los conductores y los técnicos han advertido que las ruedas con cubiertas gruesas y flexibles ofrecen el mejor medio de locomoción a través de la arena y polvo sueltos. Por tanto, la mejor solución sería neumáticos sin aire, de metal elástico u otro material flexible.

Otros requisitos para un vehículo lunar es ser liviano y compacto, para su fácil trasporte desde la Tierra a la Luna; tener chasis y ejes resistentes y flexibles, para facilitar el movimiento sobre el áspero suelo lunar; capacidad para soportar las condiciones de vacío y grandes temperaturas extremas de la Luna, y suficiente velocidad para viajar distancias apreciables en un tiempo razonable. El vehículo debe tener lugar por lo menos para dos astronautas con su equipo, enseres experimentales y muestras de roca lunar. También son vitales los sistemas de navegación y comunicación.

Los Estados Unidos dotaron de estas características a su primer vehículo, el LRV (Lunar Roving Vehicle). En su misión inicial, llevó a dos astronautas norteamericanos en varios viajes cortos de exploración a través de la superficie lunar.

Estación central de un equipo de Experimentos Apolo para la Superficie Lunar. Este equipo contenía los muchos dispositivos usados para estudiar la Luna

NASA



Cuando el LRV de aluminio fue desplegado del módulo sobre la Luna, tenía tres metros de longitud. Cada una de sus cuatro ruedas estaba impulsada por un motor eléctrico de un cuarto de caballo de fuerza. Ruedas flexibles de tejido de alambre proveyeron excelente tracción. La energía fue suministrada por baterías eléctricas. La velocidad máxima era de alrededor de 16 kilómetros por hora; el alcance máximo, de alrededor de 65 kilómetros. El LRV podía cruzar grietas de 60 centímetros de ancho y trepar pendientes de 25°. Se dirigía mediante una barra de control accionada en forma muy parecida a la de un aeroplano. Se podía virar las cuatro ruedas, para las maniobras bruscas. El vehículo llevaba dos astronautas con su equipo. Instrumentos de navegación accionados por computadoras indicaban a los astronautas su ubicación. Una cámara de televisión sobre el LRV trasmitía a la Tierra –a través de una antena de alta ganancia– vistas de la Luna.

El peso en tierra del LRV era de 200 kilogramos y en la Luna, de algo más de 34. El vehículo podía operar inclusive si se descomponía alguna de sus piezas vitales; lo que, en efecto, ocurrió al principio.

La misión descendió en la región central del norte de la cara visible de la Luna, en el Palus Putredinis ("Pantano de Putrefacción"), entre los Montes Apeninos y el Mar de las Lluvias. En esta área existe una gran incisión parecida a un cañón, denominada Fisura de Hadley, que ha intrigado por generaciones a los astrónomos. Los astronautas hicieron tres viajes en el LRV a lo largo del frente de la montaña Apenina y de la Fisura de Hadley, reuniendo muchas muestras de roca lunar. En 1972 usaron otros LRV los astronautas del Apolo 16 y del 17.

La Unión Soviética también fabricó vehículos lunares –Lunojod ("Vagabundos Lunares") 1 y 2– con diversos mecanismos adicionales; un Lunojod es fundamentalmente una máquina automatizada no tripulada. Este vehículo lunar se "maneja" por control remoto desde la Tierra. El "conductor" terrestre ve por televisión una imagen de la superficie lunar, enviada por las cámaras del Lunojod a medida que se desplaza.

El Lunojod está equipado con dispositivos automatizados para controlar y vigilar sus propios movimientos. También lleva dentro de un gran recipiente con tapa, otros instrumentos para estudiar la Luna y realizar allí experimentos científicos.

NASA

El 20 de julio de 1969, Edwin E. Aldrin, Jr., el segundo hombre en la Luna, posa al lado de la bandera de Estados Unidos, colocada en la superficie del satélite.

A fines de 1970 el Lunojod 1 se hizo alunizar en el Mar de las Lluvias, en el hemisferio boreal de la Luna; el Lunojod 2 fue depositado, a comienzos de 1973, en el cráter Le Monnier, también boreal.

ESTUDIO DE LA LUNA

El campo de la astronomía que trata la Luna en general se denomina *selenología*. El ramo de la selenología que trata las características de la superficie lunar es la *selenografía*. La selenología se remonta a la época anterior a los telescopios; empero, la selenografía comienza con éstos. En su período inicial –desde 1609 hasta el siglo XIX– los selenógrafos estuvieron abocados primordialmente a cartografiar la superficie de la Luna; un trabajo lento y penoso, que requería años de observación paciente con los ojos pegados al telescopio y realizando un dibujo minucioso.

La moderna ciencia de la *astrogeología*, que trata la estructura y caracteres de mundos sólidos (cuerpos planetarios) en el espacio exterior, tuvo su origen en la selenografía. Otros instrumentos, además de telescopios y cámaras, han venido en ayuda del selenógrafo: espectroscopios, radar y sondas lunares.

CARACTERES DE LA SUPERFICIE DE LA LUNA

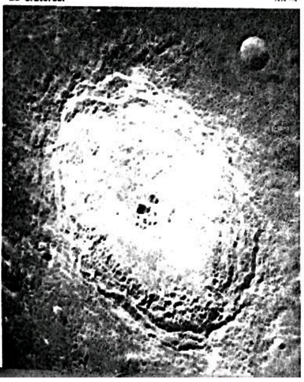
Quizá los caracteres más controvertidos de la Luna sean los cráteres; todavía es encarnizado el debate sobre su origen. En este punto, predominan dos escuelas de pensamiento: la escuela volcánica sostiene que los cráteres lunares fueron causados por erupciones de roca fundida del interior de la Luna. La escuela meteorítica afirma que se formaron por el impacto explosivo de meteoritos provenientes del espacio. Algunos astrónomos no apoyan exclusivamente ninguna, pues consideran que los cráteres fueron causados por una interacción de ambos procesos; y hay quienes piensan que por lo menos algunos cráteres lunares se deben a otros procesos geológicos, como el hundimiento o desmoronamiento de rocas, capaces de originar depresiones circulares, superficialmente semejantes a cráteres volcánicos o meteoríticos.

Algunos científicos creen que meteoros e incluso pequeños asteroides que cayeron sobre la Luna han causado la mayor parte de sus rasgos. Se ha afirmado que los impactos de grandes trozos de materia proveniente del espacio originaron la formación de los *mares lunares*; éstos tienden a mostrar forma circular y están rodeados por montañas, de modo análogo a los grandes cráteres. Los meteoristas creen que el impacto explosivo de un meteorito afectó la roca lunar en vasta escala. Luego, lava del interior de la Luna cubrió muchos kilómetros cuadrados del área, formando el mar.

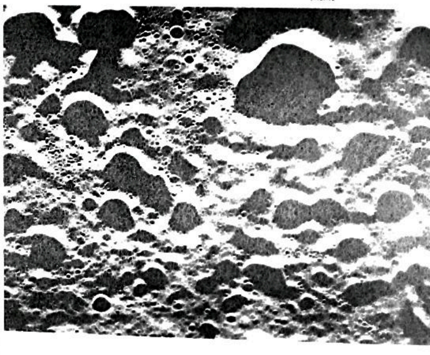
Por otro lado, los vulcanistas arguyen que los rasgos lunares surgieron por procesos propios de la Luna. Generalmente, pero no siempre, estos procesos son muy parecidos a los

Dos de los muchos miles de fotos de la superficie de la Luna tomadas por las tripulaciones de las Apolo. Izquierda: Vista desde la Apolo 11 que muestra cráteres cuyos anchos varían de 1 a 2 pies hasta 50 metros. Derecha: vista del lado oculto de la Luna, cribado de cráteres.

NASA



NASA



que moldearon los caracteres de la Tierra. Muchos científicos creen que ciertos rasgos superficiales no pudieron haber sido causados por choques al azar de meteoritos. Los vulcanólogos sostienen que los brillos peculiares y nieblas que se observan de vez en cuando en la Luna son signos de actividad volcánica. Sus oponentes afirman que son una luminiscencia de la roca lunar causada por radiación proveniente del espacio.

Las sondas lunares y las misiones Apolo de los E.U. han confirmado la existencia de algunos caracteres lunares que se asemejan notablemente a los producidos por la actividad volcánica terrestre: cúpulas y coladas de lava. Queda en pie la cuestión de si alguna vez esta actividad fue extensa e importante en la superficie lunar. La abundante fusión de rocas indudablemente se debió a impactos de grandes meteoritos sobre la Luna, los cuales, probablemente, produjeron también los grandes cráteres y los mares, o llanuras.

COMPOSICIÓN E HISTORIA

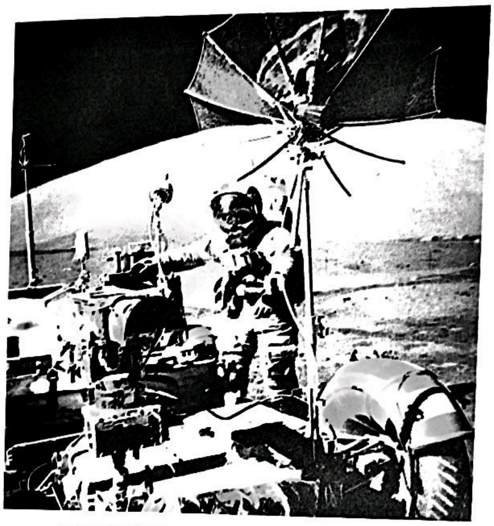
Generalmente, las rocas lunares se asemejan a las terrestres conocidas como *basaltos*, que son de grano fino, y a los *gabros*, de grano más grueso. Son éstas rocas oscuras, densas, y químicamente relacionadas, de origen ígneo.

Entre las rocas más jóvenes de la Luna están los basaltos de los mares. Estas rocas se formaron a partir de lavas oscuras que afloraron de 3.000.000.000 a 3.800.000.000 de años atrás, para formar los pisos de muchos grandes cráteres y mares.

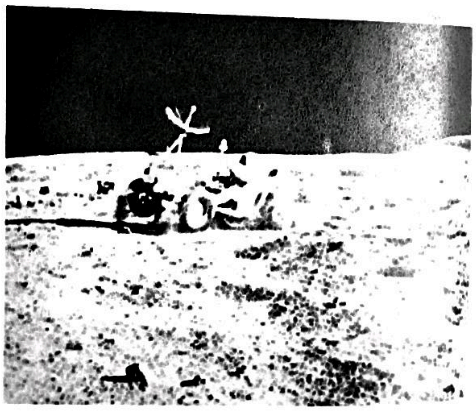
Tal como atestiguan sus rocas, la Luna ha tenido una historia larga y compleja, que comienza alrededor de 4.600.000.000 de años atrás, o antes, cuando se cree que se solidificó la corteza original.

La Luna posee una estructura complicada. Existe una *corteza* densa y rígida, de alrededor de 24 kilómetros en el lado que enfrenta a la Tierra, y de hasta 64 kilómetros de espesor en el lado oculto. Por debajo de la corteza está el *manto* de roca más densa, con composición parecida a la del gabro. Puede haber una capa parcialmente fundida a unos 965 kilómetros de profundidad.

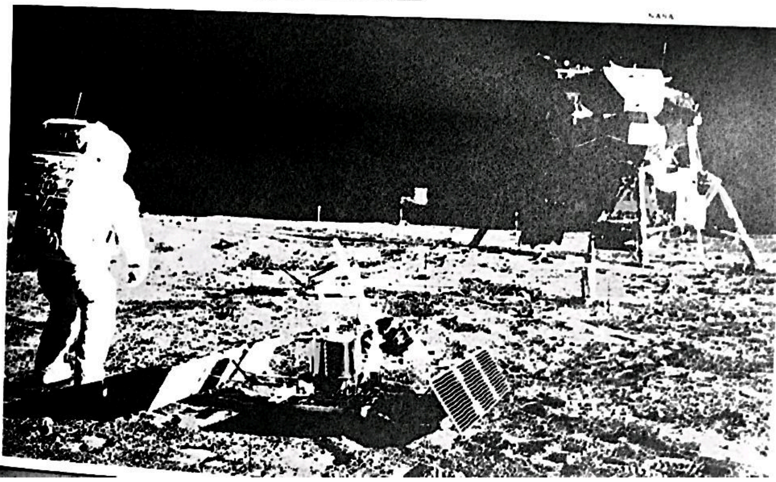
En su centro quizá haya un núcleo pesado, posiblemente líquido o semilíquido, con un diámetro de alrededor de 640 kilómetros. Su posición es excéntrica, con el centro más próximo a la cara de la Luna que enfrenta a la Tierra. Posee un campo magnético débil, cuyo origen todavía es incierto. Para un tratamiento más detallado de la Luna y de las rocas lunares, véanse los artículos "La Luna" y "Geología de la Luna", en esta misma obra.



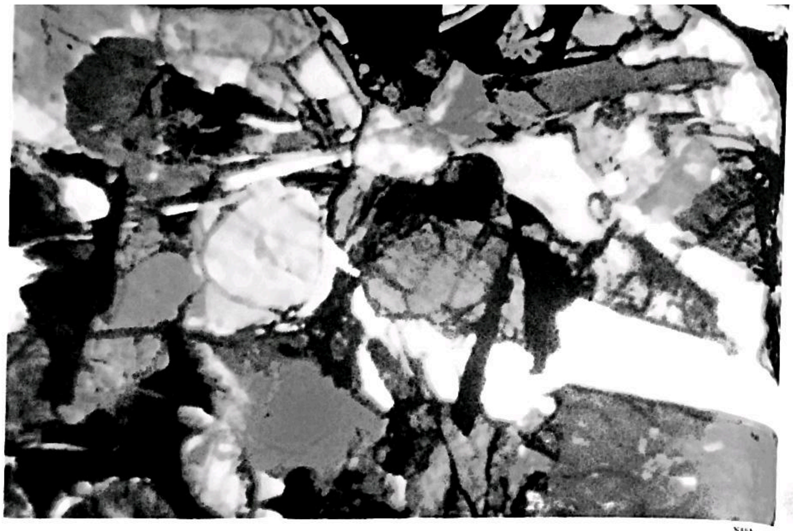
Izquierda: el astronauta Eugene Cernan sube a bordo del Lunar Rover. Abajo: Base Tranquilidad, sitio de descenso de la primera misión Apolo. El horizonte lunar está a unos 3 kilómetros de distancia.



El astronauta John Young conduce el Lunar Rover; el mismo se desplazó en la superficie de la Luna a una velocidad promedio de alrededor de 6,5 kilómetros por hora.



NASA



NASA

Microsección de una de las rocas que se trajeron a la Tierra desde la superficie de la Luna. Esta colorida roca ígnea contiene tres minerales: clinopiroxeno, plagioclasa e ilmenita.

GEOLOGÍA DE LA LUNA

Cuando el hombre dirigió por primera vez el telescopio a la Luna, hace ya más de tres siglos, vio montañas, llanuras y cráteres, y se dio cuenta de que ése era un mundo similar a la Tierra en muchos aspectos. Los astrónomos hicieron mapas de la cara lunar visible, dando nombres a sus accidentes y midiendo sus alturas y profundidades. Inmediatamente iniciaron sus especulaciones sobre el origen de estos caracteres.

A medida que la geología, o estudio de la Tierra, se desarrollaba simultáneamente con la astronomía, la naturaleza de las formas de la superficie lunar comenzó a adquirir más significación. Los científicos empezaron a comparar la Tierra y la Luna, hallando no sólo llamativas similitudes, sino también grandes diferencias entre ellas.

Con telescopios, y también con instrumentos más modernos, como el radar, los astrónomos sólo podían comprender hasta cierto punto los caracteres lunares. En la década del 60 se dispuso de mayor caudal de conocimientos, pues naves espaciales no tripuladas fotografiaron la Luna de cerca, descendieron sobre ella, tomaron muestras de sus rocas y enviaron los datos a la Tierra. En 1969, los astronautas pusieron el pie en nuestro satélite, y volvieron a la Tierra con muestras de su polvorienta superficie y sus formaciones rocosas.

Todos los datos obtenidos de la observación desde la Tierra, vuelos no tripulados a la Luna, y las exploraciones de los astronautas de las naves Apolo, se han combinado en un solo campo científico al que se denomina *geología lunar*.

GEOLOGÍAS LUNAR Y TERRESTRE

Esencialmente, la geología lunar es el estudio de la Luna por medio de sus rocas y formas del suelo. Su meta es lograr una imagen correcta y razonablemente detallada de cómo ha evolucionado. Puede sorprender oír que pronto lograremos entender más claramente el desarrollo geológico de la Luna que el de la Tierra. Esto es especialmente cierto en lo que respecta a las etapas iniciales –los primeros 2.000.000.000 de años, o algo así– de la historia de la Luna. En la Tierra, las rocas que representan este lapso generalmente han sido enterradas profundamente por otras más jóvenes.

Inclusive, en los sitios donde tales rocas terrestres muy antiguas han quedado expuestas, a menudo han sido metamorfoseadas o alteradas notablemente a causa de fuerzas y acontecimientos geológicos diferentes, hasta resultar extrañas o muy poco parecidas a su condición original. En efecto, fueron destruidas frecuentemente por erosión o fusión. En muchos casos, su presencia original sólo puede detectarse por los claros dejados en la secuencia de las rocas restantes. Aparentemente, la Luna ha sufrido muchísimo menos tales cambios. Aunque de ningún modo geológicamente "muerto", nuestro satélite, comparado con la Tierra, se halla en una especie de congelación geológica.

LA SUPERFICIE DE LA LUNA

Si bien la Luna puede parecer un caótico desierto, sus caracteres visibles pueden ordenarse y clasificarse. La superficie se divide, principalmente, entre llanuras oscuras, generalmente a nivel, llamadas mares, y las tierras elevadas, o altiplanicies brillantes y escarpadas. Tanto los mares como las tierras altas están cubiertos de estructuras circulares llamadas cráteres.

A veces, extendiéndose tanto a través de los mares como de las tierras altas por cientos de kilómetros, existen grandes fallas o grietas, que pueden ser rectas o sinuosas. Otros tipos de caracteres incluyen riscosas rugosidades en los mares, y probables formas volcánicas, como cúpulas, conos y aparentes coladas de lava. Tomaremos estos caracteres lunares por separado, comenzando por los cráteres.

Cráteres: Las formas más llamativas del terreno lunar son los cráteres, estructuras casi circulares que consisten en una pared o reborde que circunda un área generalmente a nivel. Esta área interna, o piso del cráter, puede estar por debajo del nivel general del suelo que rodea el mismo. El piso suele ser relativamente liso, o tener montañas o colinas. Los cráteres pueden parecer recientes, con rebordes elevados fuertemente definidos o escabrosos. Otros son muy bajos y casi carentes de reborde. Los de este último tipo están en apariencia muy erosionados, enterrados bajo rocas posteriores o sedimentos sueltos, o llenos de lava solidificada.

Los cráteres muy grandes se llaman circos. Buenos ejemplos son Copérnico y otro similar, Tycho, al sur. Ambos son depresiones amplias y no muy hondas, con bordes elevados y rodeadas por montones de material, evidentemente expelido de los cráteres. Éste constituye un terreno accidentado cerca de los circos, y se extiende hacia afuera, dando lugar a colinas que forman un diseño radial.

A su vez, estas colinas radiales muestran un halo de cráteres secundarios; se cree que éstos sean hoyos poco profundos excavados en la superficie lunar por la caída de fragmentos de roca expelidos de los cráteres principales. Irradiando aún más lejos de éste están los rayos: vetas brillantes en el suelo lunar. A lo largo de estos rayos se advierte en muchos lugares cadenas de cráteres secundarios. Los rayos del cráter Tycho se extienden casi hasta mitad de camino a través de la cara visible de la Luna.

Dentro de los circos Tycho y Copérnico hay terrazas concéntricas que bajan por las paredes hacia la depresión central. Probablemente estas terrazas son bloques de roca que se han desmoronado o deslizado algún tiempo después que los cráteres se formaron. En muchos lugares, algún tipo de material también parece haber rodado por los costados de éstos hacia el fondo.

En 1969, el hombre puso por primera vez el pie en otro miembro del Sistema Solar. Volvió con tesoros inapreciables: rocas, suelo, cuentas de vidrio. Desde ese momento, otros astronautas han visto la Tierra desde la superficie de la Luna. La foto de abajo muestra la Tierra como en cuarto creciente, por encima del horizonte lunar.

NASA

NASA

Roca lunar vista de cerca. A esta muestra traída por los astronautas se le atribuye una antigüedad de 4.600.000.000 de años: la edad calculada de nuestro Sistema Solar.

Los circos pueden haber tenido pisos líquidos en una época, o sea haber estado parcialmente llenos de lava, o roca fundida, lo cual parece evidente por el aspecto gelatinoso de la topografía del suelo, como ocurre en Tycho. Copérnico probablemente tuvo un piso similar, pero como es algo más viejo que Tycho, su piso ha sido parcialmente erosionado. Sobre los pisos de éstos y otros muchos cráteres existen elevaciones llamadas picos centrales, así como muchas colinas más pequeñas.

Copérnico y Tycho son ejemplos típicos de algunos de los grandes circos de la Luna. Muchos más son, evidentemente, versiones erosionadas y muy viejas de estos dos. Eratóstenes, por ejemplo, justo al este de Copérnico, se pareció una vez a este último, pero, con el paso del tiempo, impactos de pequeños meteoritos y otros procesos han modificado su topografía; ahora sus rayos son invisibles. Otros circos se asemejan a Tycho y a Copérnico en la mayoría de los aspectos, excepto que han sido llenados con algún tipo de rocas oscuras, probablemente depósitos de lava.

Ahora surge una cuestión en extremo importante: ¿qué tamaño pueden llegar a tener los cráteres? Esto depende de la definición de los mismos. En pequeños fragmentos de roca lunar se han detectado hoyos diminutos parecidos a cráteres de tamaño microscópico. También se pueden considerar muchas de las llanuras oscuras, o mares de la Luna como una clase de circos, o grandes cráteres, porque a menudo están rodeados parcialmente por montañas parecidas a murallas.

Si tan amplia variedad de formas del suelo puede denominarse "cráteres", bien podemos preguntarnos si todos se formaron por un solo tipo de proceso geológico.

Origen de los cráteres lunares. Durante siglos se ha debatido el origen de estos cráteres. El pensamiento actual se centra en dos teorías principales sobre su génesis: (1) impacto meteórico o comentario, y (2) actividad volcánica.

La teoría del impacto ha ganado rápidamente favor, en especial entre científicos de América del Norte, a partir de la década del 50. Las pruebas que la apoyan provienen de una variedad de fuentes. Primero, se comprobó que cuerpos de tamaño considerable, meteoritos o posiblemente hasta pequeños asteroides, chocaron en ciertas ocasiones con la Tierra; empero, muchos cuerpos más pequeños son destruidos por la atmósfera terrestre antes de que alcancen el suelo. La Luna carece de atmósfera, de modo que habría sufrido muchos más impactos de meteoros que la Tierra durante determinado lapso y en cada área determinada.

Segundo, existen muchos circos en la Tierra que son grandes y, aparentemente, de impacto; el Cráter Meteórico de un kilómetro y medio de diámetro, en Arizona, y la estructura Manicouagan, de 60 kilómetros de diámetro, en Canadá, son ejemplos probables. Casi cincuenta de estas formaciones muestran indicios de lo que pareciera un impacto meteorítico.

Estos indicios consisten en fragmentos de ferroníquel, un componente común de los meteoritos; fusión y recristalización de la roca causada por los efectos de la caída de un meteorito, y minerales insólitos que se forman bajo elevadas presiones y temperaturas, producidas generalmente por impacto.

Tercero, muchos cráteres lunares –incluyendo los circos Copérnico y Tycho– semejan hoyos terrestres causados por impactos de cohetes y explosiones. La proporción de profundidad a diámetro de muchos cráteres lunares es igual a la de los hoyos mencionados.

Cuarto, los asteroides, ocasionalmente, se aproximan mucho a la Tierra y a la Luna. En los miles de millones de años pasados, por lo menos unos pocos de tales cuerpos deben haber chocado con la Luna. Si Copérnico y circos similares no son las cicatrices de tales choques, ¿dónde están éstas?

Los partidarios de la teoría volcánica desestiman los argumentos que se dan arriba; señalan que el aparente origen por impacto de muchos cráteres terrestres y lunares no se ha probado realmente. Muchos circos lunares se asemejan a ciertas enormes depresiones terrestres llamadas calderas. Una caldera se forma cuando se derrumban las paredes de un cráter volcánico, agrandándolo. Sin embargo, las más grandes calderas conocidas no se acercan ni con mucho al tamaño del mayor de los circos lunares.

De todos modos, la teoría volcánica de la génesis de circos como Copérnico o Tycho o como otros similares, parece constituir en la actualidad la opinión minoritaria entre los científicos. Empero, a pesar de esto, existen cráteres lunares que incluso los teóricos impactistas han de admitir como de origen volcánico.

Las fotos de abajo muestran secciones delgadas de roca lunar. Los colores en las secciones están causados por la interacción de la luz polarizada con la estructura cristalina de los diversos minerales. Cada color que aparece en las secciones así expuestas representa, usualmente, un mineral distinto.

NASA

Éstos tienden a ser mucho más pequeños, generalmente, que los comunes por impacto. La mayoría de los cráteres en cadena (aquellos que tienden a presentarse en líneas o hileras) casi ciertamente es volcánica. Se piensa que emergen, como muchos volcanes terrestres, a lo largo de fracturas en la corteza, de donde surge, o hace erupción, roca fundida o gas.

Algunos de los cráteres pequeños pueden haberse formado por el impacto de rocas que cayeron después de haber sido expelidas en erupciones volcánicas. Éstos, originados por impactos volcánicos, serían llamados cráteres secundarios, al igual que los hallados en torno a los grandes cráteres de impacto.

Los circos guardan algún parecido con los mares circulares; a menudo estos últimos también están rodeados por terreno escabroso, no muy distinto en apariencia al de las paredes de cráteres. Como muchos circos, los mares tienen pisos oscuros de lava solidificada.

Mares. En contraste con los cráteres, los mares generalmente son mucho más grandes y cubren miles de hectáreas de la cara visible de la Luna. Como muchos cráteres lunares, probablemente fueron originados por impactos de grandes cuerpos provenientes del espacio. Un cuerpo celeste excavó al caer una cuenca; después, ésta se rellenó con erupciones de lava, posteriormente endurecida para formar el mar en sí.

Generalmente, los mares están a nivel, pero a menudo presentan crestas bajas y elevaciones en forma de cúpulas. Cantidades inmensas de cráteres más pequeños criban sus superficies. Las crestas parecen arrugas y pueden ser pliegues en la roca, o grietas a través de las cuales ha fluido lava.

Las cúpulas también pueden ser volcánicas. (El nombre cúpula se da asimismo a las colinas de las regiones más altas.) Las Colinas Mario son cúpulas clásicas. Parecen ser volcanes que representan una actividad tardía, a continuación de las principales erupciones que produjeron la roca de las regiones de mares.

Los mares generalmente son oscuros, en contraste con las zonas altas circundantes, que tienen apariencia más brillante, son más ásperas, escarpadas y se elevan más que los mares.

Tierras altas lunares. Las tierras altas lunares, o altiplanicies, constituyen gran parte de la cara visible de la Luna, y casi toda su cara

oculta. La razón de este extraño fenómeno topográfico ciertamente no se conoce hasta el presente (si bien se han propuesto algunas teorías). Las altiplanicies son ásperas y están densamente picadas, con cráteres formados dentro de otros cráteres o sobre ellos.

El número de cráteres, mayor en las altiplanicies que en los mares, parece indicar que éstas han sido bombardeadas por meteoritos, o cuerpos similares, durante mucho más tiempo; en consecuencia, como formas de suelo, las altiplanicies son probablemente más antiguas que los mares. La corteza lunar original, si todavía existe, o está expuesta en cualquier lugar, debe ser parte de las altiplanicies.

Grietas. Otras características de la superficie que contribuyen a la escabrosidad de la faz de la Luna son las ya mencionadas grietas; éstas, que pueden cruzar altiplanicies y mares, han sido tema de especulación desde que el primer telescopio examinó el satélite. Algunas presentan un trazado recto y parecen gigantescas fracturas o fallas. Enormes bloques de la corteza lunar probablemente se han deslizado hacia abajo, a lo largo de estas fracturas, de uno u otro lado.

Un tipo diferente de grieta serpentea violentamente durante muchos kilómetros, como los meandros y vueltas de un río. En efecto, esta apariencia sugirió que dichas hendiduras fueron valles de ríos en un tiempo ya remoto, cuando la Luna probablemente tenía atmósfera y agua. Sin embargo, ninguna de las muestras de roca lunar analizadas hasta ahora presenta más que un rastro ínfimo de agua, si es que se puede hablar de rastro alguno. De acuerdo con las pruebas actualmente disponibles, nuestra Luna nunca tuvo cantidad apreciable de agua en la superficie.

NASA

El Dr. Harrison H. Schmitt, geólogo civil, hizo valiosas observaciones del paisaje lunar durante la exploración de la astronave Apolo 17. Aquí se lo ve recogiendo muestras rocosas.

NASA

Fotomicrografía de un corte delgado de roca lunar tomada con luz polarizada, que revela estructuras minerales en un hermoso despliegue de colores.

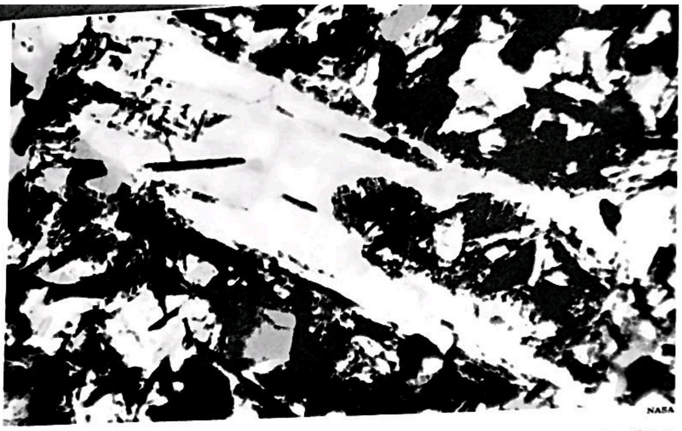
Otra posibilidad es que las grietas sinuosas sean canales para el drenaje de lava proveniente de la actividad volcánica. Aun otra, es que sean valles socavados por el flujo de ceniza volcánica, que es un material de grano fino expulsado durante las erupciones.

En este examen de la faz de la Luna, de vez en cuando hemos mencionado las rocas y otros materiales que constituyen sus características. Sólo una vez que las muestras de la superficie lunar fueron traídas a la Tierra, los científicos pudieron determinar con precisión la naturaleza de las rocas de nuestro satélite, así como la composición y estructura probables de éste.

MINERALES Y ROCAS LUNARES

Casi en todas partes, la Luna está cubierta por capas de polvo y fragmentos más grandes de roca. No se sabe con certeza el espesor exacto de estas capas en todas partes, pero probablemente sea de un promedio de varios metros. En algunas áreas puede ser mucho más profundo. El llamado polvo lunar, suelo o manto, se conoce técnicamente como regolito.

Regolito. Probablemente, varios procesos lo formaron. Éstos incluirían el impacto de cuerpos que caen y rompen la roca sólida, erupciones volcánicas que esparcen detritos en extensas áreas, deslizamientos de tierra, terremotos lunares y el efecto pulverizante de la intensa radiación proveniente del Sol y del espacio exterior.



NASA

En esta muestra, el mineral amarillo y naranja rojizo es piroxeno; el negro es ilmenita, y el azul blancuzco es plagioclasa (un tipo de feldespato).

El regolito lunar contiene variedad de materiales: fragmentos inalterados del lecho rocoso subyacente, pedazos de roca fundida y glóbulos parecidos a vidrio, ceniza volcánica y pequeñas cantidades de materia meteorítica. Las sustancias parecidas a vidrio, especialmente las esférulas, probablemente se produjeron por el impacto de meteoritos. Estos últimos, al caer, trituraron y fundieron la roca lunar, la que se volvió a solidificar, pero dando un material parecido al vidrio.

Mucho del regolito consiste en mármol brecha, depósito de fragmentos de roca con ángulos cortantes de gran tamaño, comprimido juntamente con varios otros trozos de minerales. Generalmente, el regolito presenta una consistencia floja. Aquí y allá, el material quebrado se aglutina para formar un tipo de compuestos de roca secundaria.

Sin embargo, algunas muestras de regolito difieren del promedio; contienen una elevada proporción de potasio, fósforo y metales insólitos, conocidos como tierras raras. Los científicos creen que esta porción del regolito representa los restos de una antigua corteza que una vez cubrió la superficie de la Luna, alrededor de 3.500.000.000 de años atrás. La corteza fue destruida luego por erupciones volcánicas e impactos meteoríticos.

El regolito cubre el lecho rocoso de la Luna. La naturaleza del regolito se ha inferido a partir de los pedazos más grandes de roca, que son fragmentos de este lecho rocoso.

Rocas lunares sólidas. Las rocas sólidas de la Luna, y una cierta proporción de su regolito, parecen ser de origen ígneo. Las rocas ígneas son las que se forman directamente por la solidificación de materia rocosa fundida, o magma. Los científicos han determinado que la mayoría de las rocas lunares son similares a las rocas ígneas terrestres llamadas basaltos y gabros. Los basaltos y gabros son pesados y, generalmente, de color oscuro.

Las rocas lunares también son generalmente pesadas y oscuras. Van desde material de grano fino hasta el de grano grueso. El basalto está muy finamente cristalizado, mientras que el gabro es relativamente grueso. Los granos pueden ser de tamaño igual o variado. Las muestras lunares obtenidas hasta ahora son generalmente vesiculares, es decir, llenas de cavidades. Se cree que estas cavidades fueron producidas por gas atrapado por un tiempo en la roca, mientras estaba fundida; los gases, luego, escaparon al espacio cuando la roca líquida se vertió sobre la superficie lunar. Los basaltos y gabros son ricos en minerales que contienen una elevada proporción de metales, como hierro, magnesio, aluminio, calcio y titanio.

Minerales lunares. Los elementos raramente se presentan en estado nativo, o libres, en las

rocas terrestres y lunares. A menudo se encuentran combinados en mezclas de compuestos químicos naturales, o minerales. Los tipos más comunes de compuestos minerales son los silicatos (combinaciones de silicio y oxígeno por un lado, con diversos elementos metálicos por el otro) y los óxidos (combinaciones de oxígeno con varios otros elementos). Un óxido importante es la sílice, un compuesto de silicio y oxígeno. También son significativos diversos óxidos de hierro o titanio.

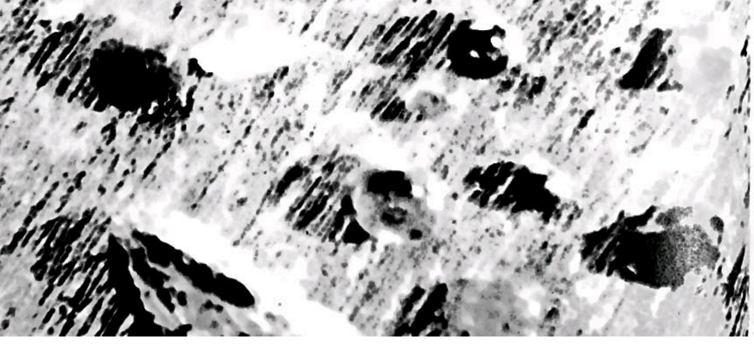
Es inmensa la variedad de minerales de silicato. Su clasificación depende de la clase de metales que contienen y la proporción de silicio y oxígeno. Los silicatos de importancia para el tema de este artículo son los feldespatos: silicatos de potasio, sodio, aluminio y calcio; los piroxenos y olivinas: silicatos de calcio, magnesio y hierro. Las olivinas tienen menor tenor de silicio y oxígeno que los piroxenos.

La mayoría de los minerales que se hallan en las rocas de la Luna son muy parecidos a los terrestres correspondientes. Varios, no obstante, son completamente desconocidos en la Tierra. Los minerales importantes de las rocas lunares son piroxenos, olivinas, feldespatos con elevado tenor de calcio, ilmenita (titanio oxidado ferrífero) y una forma volcánica de cuarzo (una forma de sílice) llamada cristobalita. Todos éstos, y otros, señalan una química lunar más rica en calcio, magnesio, aluminio, hierro, titanio y cromo que la Tierra. En cambio, las rocas de la Luna tienen relativamente menos potasio, sodio, silicio, oxígeno y agua.

Parece haber escasez de rocas graníticas en la Luna. Los granitos terrestres son ricos en minerales que contienen mucho sodio, potasio, litio y sílice. En contraste con los basaltos y gabros, los granitos son menos densos y de color más claro. En la Tierra, los granitos constituyen gran parte de los continentes, mientras que los basaltos integran principalmente la corteza de la Tierra que subyace en las diferentes cuencas oceánicas.

Historia de las rocas lunares. A partir de la naturaleza química y física de las rocas lunares, los científicos han deducido que muchas de ellas se enfriaron muy pronto; es decir, eran lavas que hicieron erupción sobre la superficie de la Luna, se congelaron rápidamente y produjeron la roca de grano muy fino llamada basalto. Dos magmas que se enfriaron más lentamente produjeron roca de grano más grueso: el gabro. Éste se halla generalmente en rocas ígneas incrustadas o metidas dentro de otras circundantes más antiguas, no dispersas sobre la superficie del suelo.

Vista microscópica de esférulas de vidrio halladas en el suelo lunar. Los científicos creen que estas cuentas se formaron bajo muy elevadas temperaturas, quizá como resultado del impacto de un meteorito.



NASA

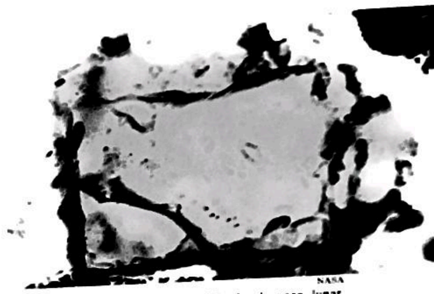
Los geólogos también han aumentado su conocimiento sobre la clase de magma del que provinieron las rocas lunares; probablemente fue muy fluido, o escurridizo, y puede haber tenido una elevada temperatura de 1.300°C. También fue un magma "seco", o de bajo tenor de agua. Este magma puede formar, o formó una vez, un depósito de líquido dentro de la Luna, o puede haber sido producido por la fusión local de rocas lunares a cierta profundidad por debajo de la superficie. Las rocas lunares también indican algo sobre el ambiente pasado y presente de la Luna, como la presencia de un antiguo campo magnético y la intensidad anterior de los rayos cósmicos.

Ambiente lunar. El agua, casi inexistente en la mayor parte de las muestras de rocas, parece demostrar que la Luna nunca tuvo una provisión apreciable de ésta, especialmente en la superficie. El hecho de que se haya descubierto hierro libre en varias muestras de roca demuestra que oxígeno y agua atmosféricos casi no existieron en ella. El hierro es rápidamente atacado por agua y oxígeno, para formar hematita (herrumbre) y otros minerales.

Varias rocas lunares muestran pruebas de choque o impacto. Han sido golpeadas repetidamente por cuerpos probablemente espaciales, de modo que algunas partes se fundieron y se solidificaron, dando una materia parecida al vidrio.

Se halló que algunas de las muestras de roca sólida y regolito contienen gases de varias clases. Estos gases probablemente provinieron de dos fuentes principales: desintegración radiactiva de ciertos elementos minerales, como uranio y torio, y partículas de radiación llegadas desde el espacio exterior. En el último caso, las partículas penetraron la superficie lunar y quedaron atrapadas ahí. En efecto, son materia y energía provenientes del Sol y las estrellas, y no partes de la Luna.

Densidades y distribución de las rocas lunares. En algunos de los mares existe una concentración de masa, o materia densa bajo la superficie, a la cual se ha dado el nombre de *mascon*. Los mascons fueron detectados porque los vehículos que describían su órbita en torno a la Luna eran atraídos hacia abajo ligeramente, en dirección a esos mares. Los científicos todavía no pueden explicar qué son los mascons, aunque tienen algunas teorías. Por otro lado, por debajo de varios cráteres, las densidades de roca son menores que en el terreno circundante; los vehículos que vuelan sobre estos cráteres se elevan ligeramente en sus órbitas.



NASA

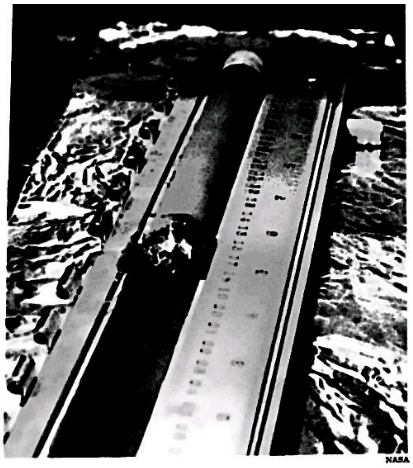
Fotomicrografía de una sección delgada de roca lunar, para uso en la identificación de minerales.

Los astronautas, al explorar los mares y montañas lunares, hallaron rocas de las altiplanicies en éstos. Fueron arrojadas allí por el impacto de enormes meteoritos que cayeron en las altiplanicies. Estas rocas son muy diferentes en composición a las rocas de mar y, en efecto, semejan notablemente ciertas rocas extrañas de diferentes zonas de la Tierra.

Estas insólitas rocas terrestres y las lunares correspondientes, se denominan anortositas, pues consisten casi enteramente en un feldespato rico en calcio, llamado anortita. Otros minerales y sus elementos están presentes en cantidades menores, notablemente la ilmenita (titanio oxidado ferrífero), y minerales que contienen cromo y aluminio. La anortosita se relaciona químicamente con el gabro. En la Tierra existe una variedad de anortositas ampliamente diseminada en la corteza.

Todas las anortositas terrestres conocidas son muy antiguas: la más vieja hallada hasta el presente se calcula que tiene más de 3.000.000.000 de años, edad comparable con la de las rocas lunares. Las anortositas son menos densas que los basaltos y gabros. El magma del que derivó la anortosita debe haber sido muy caliente y fluido.

Estos hechos sugieren cierta similitud entre la historia de la Luna y el desarrollo temprano de la Tierra. ¿Es posible que la corteza original de ésta fuese anortosítica? ¿Formó la anor-



NASA

Sedimento lunar no perturbado, fotografiado al abrir el tubo de testigos. El material se asemeja a la arena aluvial, con fragmentos de roca esparcidos y partículas vidriosas.

Examinando una roca lunar en el Laboratorio de Recepción Lunar del Centro de Naves Espaciales Tripuladas en Houston, Texas. La estructura y composición de cada muestra se analizaron cuidadosamente.

tosita, en algún momento, las masas continentales de la Tierra, en vez del granito? Si es así, las anortositas terrestres desaparecieron en su mayoría, siendo reemplazadas, o absorbidas, por efusiones posteriores de basalto y granito.

Las rocas de las altiplanicies lunares son diferentes de las rocas de los mares; contienen muchos más minerales, y también son distintas en composición del material de las tierras bajas. Su estructura es muy compleja; consiste en brechas antiguas compuestas por aglutinaciones aún anteriores de fragmentos de roca; evidentemente, experimentaron un desarrollo complicado. Todavía está por verse si las rocas recogidas por los astronautas de las naves Apolo en las altiplanicies son típicas de todas las rocas de las montañas lunares.

EDAD DE LA LUNA

En general, los accidentes superficiales de la Luna, y sus rocas, son muy antiguos, del orden de varios miles de millones de años. Cómo se deducen tales cifras para fechar las rocas y formaciones geológicas es muy interesante.

Primero, a los geólogos les interesan las edades relativas de los caracteres; es decir, cuál es más antigua o más reciente que cuál. En general, cuanto más vieja sea una roca o formación del suelo, tanto más desgastada o erosionada tiende a ser. También puede estar química y físicamente alterada. Las rocas y accidentes más nuevos tienden a descansar sobre los más antiguos, pero éste no es siempre el caso: por ejemplo, una masa de roca fundida puede abrirse camino hacia arriba, por entre rocas más antiguas y, de este modo, llegar a yacer debajo de las superficiales. El estudio de las secuencias de las capas rocosas, o estratos, así como el de las otras clases de masas rocosas, se conoce como *estratigrafía.*

Segundo, los geólogos desean determinar la edad *absoluta* de una roca, o sea los años que cuenta; esto se hace usualmente por mediciones químicas y de radiactividad de los tipos de elementos que contiene.

La cantidad de tiempo que una roca, independientemente de su edad exacta, haya sido expuesta a la radiación, especialmente la proveniente del Sol o del espacio exterior, también puede determinarse a partir de su composición química y radiactividad. De esta manera, los científicos hallaron por cuánto tiempo estuvieron expuestos a las condiciones del espacio los meteoritos y las rocas de la superficie lunar. Esta es una clave valiosa para su historia.

Datación de las rocas lunares. Según los patrones terrestres, éstas son muy antiguas. Las más recientes de las conocidas superan los 3.000.000.000 de años, y otras los 4.000.000.000. Las lavas del Mar de la Tranquilidad cristalizaron hace unos 3.700.000.000 de años; las del Océano de las Tempestades, unos pocos cientos de millones de años después. El Mar de las Lluvias se formó alrededor de 3.900.000.000 de años atrás, y es uno de los accidentes lunares más antiguos.

Aunque suene extraño, los pequeños fragmentos del regolito propiamente dicho son realmente más antiguos que la roca sólida al parecer subyacente de los mares explorados hasta ahora. Llegan hasta los 4.600.000.000 de años. He aquí, pues, una clase de secuencia estratigráfica invertida, en la que los depósitos superiores son más antiguos que los inferiores. El regolito donde se asentó la nave Apolo ha estado sobre la superficie lunar a merced de la radiación solar y espacial por períodos de un millón a varios cientos de millones de años.

Este problema estratigráfico de la corteza lunar es desconcertante, y ha llevado a los geólogos a considerar varias teorías nuevas, a fin de explicarlo.

Estratigrafía lunar. Para resumir la situación en la superficie de la Luna, tenemos la presunta roca sólida de la corteza a alguna profundidad por debajo del regolito. Por lo menos parte del último debe de haber provenido del lecho rocoso, por impacto meteorítico, radiación y otras fuerzas erosivas.

Para explicar por qué el regolito generalmente es más antiguo que el lecho rocoso, los geólogos han considerado varias teorías alternativas: una, considera que el regolito de cualquier punto de la superficie es una mezcla representativa de porciones de la corteza lunar. Fragmentos de amplias áreas de la Luna, incluyendo las muy antiguas altiplanicies, pueden haber sido diseminados en grandes extensiones por fuertes y continuos impactos meteoríticos y erupciones volcánicas; así, la edad del regolito en cualquier lugar puede representar la edad de la corteza lunar en su totalidad. Otra

posibilidad es que el lecho rocoso haya sido empujado, en forma de magma, dentro del regolito ya existente, más antiguo que el mencionado lecho.

Es obvio que la edad de las rocas lunares consideradas en su totalidad es muy grande y que la Luna, como cuerpo celeste es aún más vieja: debe contar alrededor de 5.000.000.000 de años. Las rocas más antiguas descubiertas hasta ahora en la Tierra tienen alrededor de 3.500.000.000 de años; esto no significa necesariamente que la Tierra sea más joven que la Luna; lo probable es que sus rocas más viejas aún no hayan sido halladas, hayan desaparecido o se encuentren tan alteradas que no se puedan medir con precisión.

Combinemos ahora estas piezas del rompecabezas del conocimiento –topografía, rocas y edades– para obtener alguna clase de imagen integral de la probable historia evolutiva de la Luna.

EVOLUCIÓN DE LA LUNA

Formación de la Luna. Por supuesto, el desarrollo de la Luna comenzó con su formación, alrededor de la cual existe, hasta el presente, poco acuerdo entre los científicos. Las teorías principales sostienen que la Luna nació de la Tierra en alguna etapa temprana, o quizá relativamente tardía, de la historia de esta última; o que la Luna y la Tierra se formaron independientemente de materiales casi idénticos en los tiempos en que se originó todo el Sistema Solar; o que la Luna fue capturada por la Tierra. Según estas teorías, ambas pueden tener la misma edad, o bien ser una más joven que la otra.

Una teoría, popular en un momento dado, decía que la Luna se separó de la Tierra, la que una vez tuvo más masa que ahora. George Darwin, científico del siglo XIX, supuso que esto ocurrió cuando la Tierra primigenia se hallaba en estado de fusión. Recientemente, la idea ha sido revivida por algunas autoridades.

Una teoría ampliamente apoyada hoy en día dice que la Luna se formó de la misma nube de gas y polvo que la Tierra. Esto ocurrió al mismo tiempo o, probablemente, después de la formación de la Tierra. A pesar de que aparece tanto hierro en las rocas de la superficie de la Luna, su densidad total es baja, comparada con la terrestre: alrededor del 60 % la de esta última. En otras palabras, buena parte de la Luna, debajo de su superficie, tiene que ser considerablemente menos densa que las rocas corticales. Esto puede explicarse si se considera que la Luna, en su totalidad, es deficiente en hierro. Por otro lado, se cree que la Tierra posee en su centro un núcleo metálico pesado, muy rico en hierro. Esta situación explica su mayor densidad, si bien muchas de sus corticales no son tan ricas en hierro como las rocas lunares correspondientes.

Para explicar esta teoría, varios científicos suponen que la Luna se formó después del núcleo de hierro de la Tierra. Esto pudo ocurrir de varias maneras. Una de ellas sería que los minerales de silicato, evaporados por el intenso calor de la Tierra primitiva, constituyeran nubes o anillos de sedimento, o de materia rocosa, los cuales quedarían suspendidos en la densa atmósfera de ese período. La Luna habría sido formada por la condensación y acumulación de tales nubes o anillos.

Desarrollo inicial de la Luna. Durante la génesis de la Luna, o poco después, ésta probablemente se volvió muy caliente. Ello fue causado por la compresión de los materiales que la formaban, por el calor producido en la desintegración de los elementos radiactivos de

El astronauta Edwin Aldrin tomando una muestra de regolito. La tripulación del Apolo 11 recogió cantidad de rocas y suelo para estudio. Con estas piezas se inició la verdadera investigación geológica lunar.

vida breve, y por el calor proveniente de otras fuentes. Las altas temperaturas pueden haber ocasionado una amplia fusión y refusión de las rocas lunares, suficientes para que pudieran aparecer diferentes clases de rocas ígneas. Esto puede explicar la posible diferencia entre las planicies basálticas (mares) y las altiplanicies anortosíticas o, posiblemente, graníticas.

Mientras esto ocurría, cayó sobre la Luna una gran cantidad de cuerpos llamados planetesimales, provenientes del espacio, los cuales aportaron sus elementos y formaron los cráteres más antiguos o muy antiguos, en las áreas de altiplanicies y en la cara oculta. La caída de estos cuerpos puede considerarse como la última etapa del primer período de desarrollo de la Luna.

Segunda etapa de desarrollo de la Luna. Durante este tiempo, varios cuerpos, asteroides o pequeñas lunas, crecían en la vecindad de la Luna. Ésta los atrajo, y sufrió así impactos catastróficos, que formaron las cuencas llenadas luego por los mares circulares: los de la Serenidad, de las Crisis, de los Humores, de las Lluvias, Oriental y otros. Estos tremendos impactos causaron algo de fusión en las rocas lunares, pero no produjeron realmente las efusiones de lava en gran escala que llenaron las cuencas y luego formaron los mares. Muchos cráteres de impacto, tales como Arquímedes y Platón, se produjeron después de la excavación de las cuencas de los mares, pero antes de que estas cuencas se rellenaran.

Todavía surgieron más rasgos geológicos antes de las coladas de lava de los mares. La actividad volcánica en las altiplanicies lunares (o en lo que ahora son las altiplanicies) llenó cráteres como Tolomeo y Alfonso, con lava y ceniza. Este material se asemeja a las lavas y cenizas de los mares posteriores, pero tiene más cráteres a causa de su mayor edad. Las lavas de las altiplanicies también son más vidriosas, por el número mayor de impactos que han sufrido, y que les ocasionaron la refusión.

Ahora llegamos a un acontecimiento en la historia lunar que realmente ha sido fechado por el hombre: las erupciones de lava que produjeron el Mar de la Tranquilidad, donde se fijó el lugar de descenso de la astronave Apolo 11. Esta actividad ígnea ocurrió 3.700.000.000 de años atrás, que es el tiempo trascurrido desde la cristalización de las rocas del mar, a partir de la lava existente. Probablemente, esto fue acompañado por coladas similares en otras cuencas de mares, las que así se llenaron. Otros mares probablemente se formaron en épocas posteriores. Erupciones de lava renovadas de tiempo en tiempo cubrieron las anteriores coladas. Podemos ver áreas como ésta en el Mar de la Serenidad, donde las lavas más recientes son notablemente más oscuras que las de coladas anteriores.

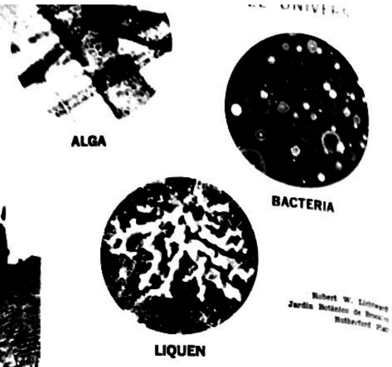
Último período de la historia lunar. Alrededor de 3.000.000.000 de años atrás, la faz de la Luna alcanzó, esencialmente, su apariencia actual. Sin embargo, la evolución geológica lunar continuó en forma de actividad volcánica ocasional en los mares y altiplanicies. Los volcanes de las Colinas Mario son de este período relativamente tardío, así como los cráteres en cadena, las grietas sinuosas y las regiones de altiplanicies más oscuras.

Empero, los eventos más espectaculares de esta época fueron los impactos que produjeron circos, como el de Copérnico. Los agentes fueron probablemente enormes meteoritos, o cometas que se desplazaban a velocidades tremendas. Cuando uno de estos cuerpos se estrelló contra la superficie lunar, produjo una enorme nube de escombros, de la que se dispararon largos chorros en todas direcciones, por cientos de kilómetros. A medida que estos materiales se asentaron en la superficie, formaron los sistemas de rayos de grandes cráteres, entre ellos Copérnico, Aristarco, Tycho y otros. Los extraños puntos rojos que se han visto en Aristarco pueden ser actividad volcánica subsistente aún después del impacto inicial.

Mientras tanto, una lluvia lenta y despareja de material meteorítico menor sigue cayendo sobre la Luna. El sismómetro (detector de temblores) dejado en el Océano de las Tempestades registró los impactos como sucesos ocasionales. Esta lluvia ha causado los incontables cráteres pequeños, y hasta microscópicos, que criban las rocas y el manto de la Luna. Por supuesto, el manto, o regolito, es en gran medida el resultado de este proceso.

Mucho se ha aprendido sobre la Luna en sólo unos pocos años. Pero los descubrimientos traen nuevos misterios. Ese globo con cicatrices que ha circundado la Tierra por tan larguísimo tiempo, será el blanco de la investigación en los años venideros.

Archivo Bettman



Robert W. Lichtwardt
Jardín Botánico de Brooklyn
Rutherford Platt

Izquierda: ilustración de invasores marcianos para la novela de H. G. Well *La guerra de los mundos*. Arriba: algas, bacterias y líquenes terrestres, simples formas orgánicas que pueden asemejarse a posibles formas de vida en Marte.

VIDA EN OTROS MUNDOS

"Dos grandes ojos oscuros me observaban fijamente. La masa que los contenía, o la cabeza, era redondeada, y podría decirse que tenía una cara. Había una boca bajo los ojos, cuyos bordes sin labios temblaban y palpitaban echando saliva."

Así describió a un marciano el destacado escritor inglés H. G. Wells en su famosa novela *La guerra de los mundos*, publicada en 1898. Relataba cómo los marcianos, invasores de la Tierra, monstruos como pulpos, con una inteligencia muy superior a la del hombre, casi destruyeron la civilización humana.

El relato de Wells, pese a su fantasía, aún refleja el pensamiento de muchos acerca de posible seres vivientes en otros mundos. El hombre consideró durante largo tiempo las estrellas, el Sol, la Luna y los planetas como hogares de dioses y demonios. En los innumerables mitos, el hombre aparece como la víctima de estas criaturas, o ayudado y civilizado por ellas. En algunos cuentos, los seres extraterrestres son completamente humanos, al menos físicamente. Estos temas sobreviven en la ficción científica y en historias fantásticas, incluyendo múltiples informes de platillos voladores. Mucha gente cree que existe vida bien desarrollada en alguna parte del Universo, fuera de la Tierra. Creen que algunas de estas posibles formas vivientes pueden ser tan o más inteligentes que el hombre.

En cambio, muchos, incluyendo científicos, consideran que los mundos allende la Tierra carecen de vida; o que, a lo sumo, sólo poseen formas de vida más inferiores como bacterias, algas y líquenes. Estas ideas son más correctas que las anteriores, al menos en lo referente a nuestro Sistema Solar.

El ramo de la biología sobre la posibilidad de vida en otros mundos se denomina exo*biología*, o sea "biología exterior [a la Tierra]". La exobiología se ocupa, además, del interrogante fundamental sobre el origen de la vida. Aquí, examinamos las probabilidades que exista más allá de la Tierra y describimos intentos del hombre para comunicarse con civilizaciones extraterrestres. Primero, sin embargo, consideremos los fundamentos químicos de la vida y cómo podría originarse en un planeta.

LA QUIMICA DE LA VIDA

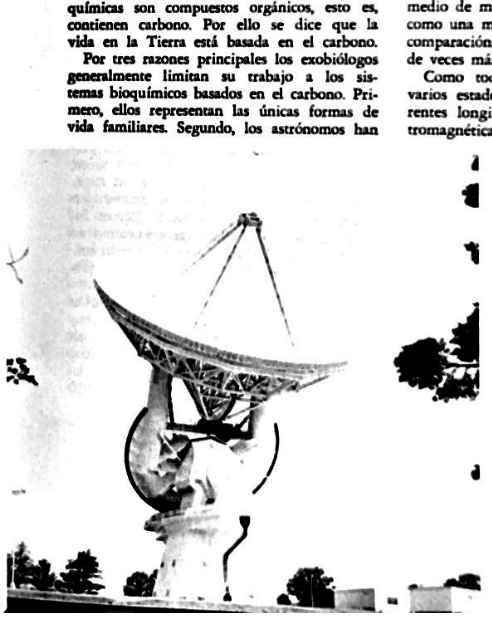
La única vida conocida hasta ahora es la que existe sobre la Tierra, que es sumamente variada e incluye bacterias y protozoos, tan pequeños que difícilmente pueden verse sin microscopio; millones de especies de insectos, gigantescos árboles secoyas y ballenas y, finalmente, el hombre mismo. Los cuerpos de estos organismos se componen de elementos químicos. Hay relativamente pocos en la materia viviente: carbono, hidrógeno, oxígeno, nitrógeno, azufre, fósforo y algunos otros, incluyendo ciertos metales; todos elementos usualmente combinados en compuestos biológicos o sustancias bioquímicas. Ciertas sustancias bioquímicas, como las proteínas, contienen todos o la mayoría de los elementos enunciados anteriormente. Otras moléculas relacionadas con la vida, como agua y bióxido de carbono, contienen sólo unos pocos. Muchas sustancias bioquímicas son compuestos orgánicos, esto es, contienen carbono. Por ello se dice que la vida en la Tierra está basada en el carbono.

Por tres razones principales los exobiólogos generalmente limitan su trabajo a los sistemas bioquímicos basados en el carbono. Primero, ellos representan las únicas formas de vida familiares. Segundo, los astrónomos han detectado compuestos orgánicos simples de tipo terrestre en el espacio exterior. Por este descubrimiento muchos exobiólogos dedujeron que las bases de la química de la vida son las mismas en todo el Universo. Tercero, estos elementos químicos se vuelven parte de los planetas cuando éstos se forman a partir de elementos en una nebulosa, o nube de polvo y gas en el espacio. Los elementos están presentes en proporción a sus respectivas abundancias en el espacio: hidrógeno, helio, carbono, oxígeno, nitrógeno y demás, con un enorme exceso de hidrógeno. En la atmósfera de un planeta primitivo, los compuestos como metano, agua y amoniaco se forman de muchos de esos elementos.

MATERIA EN EL ESPACIO EXTERIOR

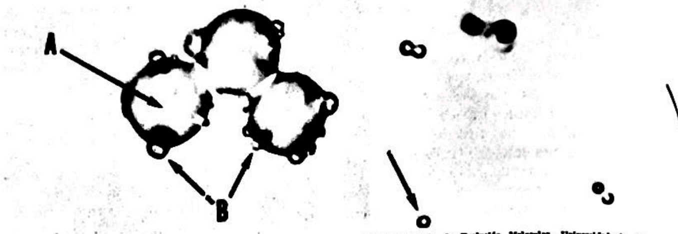
El espacio exterior contiene materia heterogénea, cuyo tamaño va desde partículas subatómicas hasta polvo cósmico y cuerpos más grandes. Pero para los criterios terrestres, el espacio exterior es un vacío. Su densidad promedio de materia es muy baja, quizá tan baja como una molécula por centímetro cúbico. En comparación, la atmósfera terrestre es millones de veces más densa.

Como toda materia, la cósmica existe en varios estados de energía. Luego, emite diferentes longitudes de onda de radiación electromagnética. Si las ondas electromagnéticas



Un gran radiotelescopio: este tipo de instrumento recibe las radioondas naturales emitidas por materia cósmica y otros cuerpos en el espacio exterior. Los radiotelescopios también se están usando para tratar de detectar posibles mensajes, o señales, electrónicos de civilizaciones avanzadas que podrían existir en cualquier otra parte del Universo.

Observatorio Nacional Radioastronómico



todas las fotos de esta página y siguientes: Instituto de Evolución Molecular, Universidad de Miami

tienen frecuencia de radio, los radiotelescopios terrestres las captarán y los científicos las analizarán para identificar la naturaleza química de los átomos y moléculas que las emitieron.

Entre estas moléculas cósmicas, los científicos ya han identificado un número de sustancias químicas relacionadas con la vida en la Tierra, o sea, idénticas a los elementos y compuestos que forman parte de los sistemas vivientes; que contribuyen a que exista la vida, o que son producidos por organismos. Las sustancias descubiertas en el espacio incluyen agua, amoniaco, metano, formaldehído, cianógeno, hidrógeno y alcohol metílico. Estos compuestos, excepto el amoniaco y el agua (excluyendo también el elemento hidrógeno), contienen carbono.

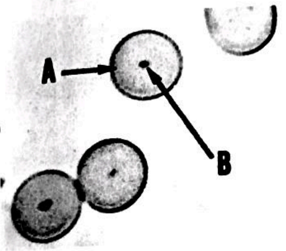
¿Cómo se originaron estos elementos químicos? Casi todos los científicos creen que esas sustancias son de origen inorgánico. Es decir, que evolucionaron a partir de sistemas químicos sin vida. Esta teoría se denomina evolución química de la vida. Está estrechamente relacionada con el origen y desarrollo de los planetas y el origen de la vida, pues al parecer es el mismo tipo de proceso químico que, se cree, ocurre en la historia primitiva de un planeta, antes que aparezcan en él seres vivientes. Algunos científicos se preguntan si esos compuestos biológicos conducen efectivamente a la formación de tales seres.

LA QUIMICA CAMBIANTE DE LOS PLANETAS

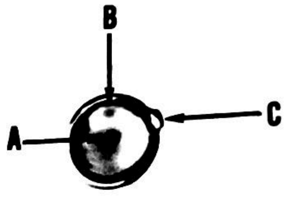
La evolución química es aún anterior a la aparición de las estrellas y planetas en el espacio. Los compuestos orgánicos simples se forman de elementos en la materia cósmica. Lentamente, la materia cósmica de toda clase se une para formar estrellas y planetas. Conforme esto ocurre, muchos compuestos orgánicos se convierten en parte de los planetas. Por ello, en esta etapa temprana de desarrollo, su composición es muy similar a la materia espacial de la cual se formaron. Luego, la química de los planetas cambia al desarrollarse a través de las edades.

Muchos científicos creen que un planeta primitivo típico, justo después de su formación, tiene una atmósfera con grandes cantidades de hidrógeno, metano, amoniaco y posiblemente vapor de agua y oxígeno. Si está lejos del Sol, alrededor del cual gira, es frío. Luego, muchos de sus gases atmosféricos pueden condensarse en nubes de gotitas o nieve; algunos gases llegan a congelarse en capas sobre su superficie. Los grandes planetas exteriores de nuestro Sistema Solar –Júpiter, Saturno, Urano y Neptuno– se encuentran en tales condiciones. Muchos científicos están convencidos de que ellos contienen, y están produciendo, sustancias químicas biológicas.

En comparación, la atmósfera terrestre es rica en nitrógeno y oxígeno. Mantiene muy bien la vida, como sabemos. ¿Se asemejó alguna vez nuestra Tierra a Júpiter o Saturno? Algunos científicos creen que sí. Hace aproximadamente 4.000.000.000 ó 5.000.000.000 de años, dicen, tenía una atmósfera de hidrógeno, amoniaco y metano, donde se desarrollaron sistemas vivientes. Otros discrepan: piensan que por lo menos algo de oxígeno libre se hallaba también en la atmósfera primitiva de la Tierra.



Formación de sustancias semejantes a proteínas (proteinoides) en esferas (microesferas) superficialmente parecidas a células vivientes. Fotos en la página anterior: izquierda, microesferas (A) formando brotes (B); derecha, brotes liberados de las esferas principales por el calor. Fotos en esta página: izquierda, sustancia proteinoide adicional (A) se acumula sobre los brotes (B), formando nuevas esferas; derecha, una nueva microesfera (B) producida del brote (A), echa un nuevo brote (C).



Si los planetas comienzan con atmósferas de hidrógeno, metano y amoniaco, ¿significa que siempre se formarán compuestos bioquímicos? ¿Originarán en cada caso estos compuestos vida en un planeta? Experimentos de laboratorio intentaron responder a tales interrogantes.

EXPERIMENTOS QUIMICOS

Un bioquímico ruso, Alexander I. Oparin, sugirió en la década del 1920 que compuestos orgánicos podrían originarse en una atmósfera planetaria de hidrógeno, metano y amoniaco. Este hecho sucedería espontáneamente si hubiera suficiente energía presente.

A partir de la década del 1950, los científicos realizaron experimentos que confirmaron la teoría de Oparin. Prepararon varias mezclas de hidrógeno, metano, amoniaco, vapor de agua y otros gases. (Una mezcla no debía contener necesariamente todas estas sustancias.) Las mezclas fueron expuestas luego a distintas cantidades de radiación, calor o energía eléctrica.

Como consecuencia, se obtuvo sorprendente número de compuestos simples orgánicos y biológicos, incluyendo aminoácidos, que son la base constitutiva de las proteínas. Algunos científicos también obtuvieron productos químicos contenidos en los ácidos nucleicos (ADN y ARN), los principales reguladores de las actividades en células vivientes.

Algunos investigadores incluso produjeron moléculas simples semejantes a las de las proteínas, denominadas *proteinoides*. Estos proteinoides estaban organizados en glóbulos minúsculos o *microesferas*, las cuales se asemejan a ciertas bacterias simples y realizan algunas funciones de células vivientes. Ninguno de los experimentadores sostiene que los proteinoides y las microesferas estén vivos. Pero consideran que estas estructuras podrían evolucionar hacia la materia viviente (protoplasma) y células, con tiempo y condiciones ambientales adecuados.

En resumen, intentemos visualizar la Tierra 4.000.000.000 a 5.000.000.000 de años atrás, como un planeta envuelto en una atmósfera nebulosa de amoniaco, metano, hidrógeno, vapor de agua y posiblemente otros gases, incluyendo oxígeno. Al actuar la energía solar, y quizá la de la misma Tierra, sobre la atmósfera, se producían más compuestos orgánicos y biológicos.

Gradualmente, estos compuestos cayeron desde la atmósfera sobre la superficie terrestre, donde formaron lagos y lagunas de una densa "sopa" químico-orgánica. La síntesis de los compuestos probablemente continuó allí. Luego, en millones de años, las sustancias bioquímicas se convirtieron en células vivientes.

Más recientemente, hace unos 3.000.000.000 ó 4.000.000.000 de años, la atmósfera de la Tierra comenzó a cambiar; se volvió rica en nitrógeno y oxígeno. Algunos científicos relacionan este cambio con la aparición de las plantas verdes, que liberaban oxígeno en la atmósfera. También piensan que la compleja vida actual de la Tierra pudo evolucionar sólo en un ambiente oxigenado, pese a que sabemos que existen ciertos tipos de bacterias en condiciones anaeróbicas o carentes de oxígeno. En contraposición, las autoridades no creen que los elementos bioquímicos puedan desarrollarse a partir de materia inorgánica en el medio ambiente rico en oxígeno propio de

la Tierra hoy en día, al menos no en escala tan grande como hace 4.000.000.000 ó 5.000.000.000 de años.

MATERIA ORGÁNICA EN METEORITOS

Los meteoritos, por ser trozos de materia del espacio exterior, interesaron a los exobiólogos desde hace mucho tiempo. Aquí, también, hubo polémica. Los compuestos químicos del carbono, detectados en unos pocos meteoritos, se desecharon generalmente como contaminación terrestre. En algunos meteoritos también se observaron estructuras peculiares, parecidas a microorganismos terrestres, como bacterias y algas. Muchos científicos consideraron que esas estructuras eran características completamente inorgánicas. Para otros la mayor parte de estos hallazgos era sugerencia de vida no terrestre.

La discrepancia ha sido, en alguna medida, superada por los científicos de la Administración Nacional Aeronáutica y Espacial (NASA) en los Estados Unidos. Ellos analizaron dos meteoritos con instrumentos sensibles de la era espacial. Cada una de las piedras (una cayó en Kentucky y la otra en Australia) rindió 18 aminoácidos, y otros compuestos de significación biológica. Muchos de los aminoácidos se encuentran en los sistemas vivientes en la Tierra, pero otros aminoácidos meteoríticos son distintos a cualesquiera de los encontrados sobre ella. Entonces, los científicos de la NASA llegaron a la conclusión de que los compuestos biológicos en los dos meteoritos se originaron en cualquier otro lugar del Universo.

Probablemente, estos compuestos no fueron producidos por seres vivientes, pero indican efectivamente que la evolución química está ocurriendo en distintas áreas del Universo; las sustancias bioquímicas van formándose a partir de materia inorgánica, y es posible que estén conduciendo a la creación de seres vivientes.

LOS PLANETAS APTOS PARA LA VIDA

¿Cuáles son las posibilidades de que la evolución química lleve a la vida en un planeta? ¿Cuántos planetas en el Universo son adecuados para seres vivientes? Los astrónomos estiman que hay 10^{20} (1 seguido de 20 ceros) estrellas en el Universo. La mayoría, pro-

NASA

Un científico en el Centro de Investigaciones Ames muestra un meteorito en el que se descubrieron aminoácidos extraterrestres, las bases constitutivas de las proteínas.

bablemente tienen planetas que giran en torno de ellas, lo mismo que nuestro Sol. (Hasta ahora, no hemos podido ver ningún planeta fuera de nuestro Sistema Solar.) Los astrónomos estiman que la vida podría haber surgido en unos 100.000.000.000 de planetas. De éstos muchos se asemejan a la Tierra y admitirían nuestro tipo de vida.

Sabemos que la vida es muy adaptable. Casi todo lugar en la Tierra es hogar de algún organismo. Ciertas bacterias y algas sobreviven a temperaturas próximas al punto de ebullición del agua; otras soportan temperaturas muy por debajo del congelamiento. Por tanto, los planetas que consideramos hostiles a la vida pueden albergar seres vivientes de algún tipo, no importa cuán simples sean.

El descubrimiento de vida de cualquier clase en otros mundos haría más que satisfacer nuestra curiosidad. También arrojaría luz sobre la evolución de la Tierra y de los organismos que la habitan. Si se están formando ahora planetas y sistemas de vida en otras partes del Universo, muchos de ellos pueden hallarse en la fase que atravesó la Tierra varios miles de millones de años atrás. Otros planetas y sistemas de vida podrían ser más antiguos que la Tierra y darnos alguna indicación acerca de nuestro futuro.

A medida que desarrollemos nuestra habilidad e instrumentos para investigar minuciosamente otros planetas, obtendremos importantes claves sobre el pasado y el futuro de la Tierra. Aun los mundos sin vida pueden decirnos algo. Si, por ejemplo, se encontrara otro planeta como la Tierra, pero carente de vida, acarrearía serias dudas sobre las teorías actuales del origen químico y evolución de ésta.

Actualmente, sólo tenemos conocimiento directo de los planetas de nuestro Sistema Solar. Éstos han sido estudiados atentamente por bastante más de tres siglos, desde la invención del telescopio. Recientemente, Venus y Marte también han sido investigados por medio de sondas espaciales no tripuladas. Este tipo de exploración continúa y se extenderá a otros planetas. ¿Descubrirán los científicos vida en alguno de ellos?

MERCURIO Y VENUS COMO HABITATS

Mercurio, el planeta más cercano al Sol, es algo más grande que nuestra Luna. Como ésta, tiene escasa o ninguna atmósfera. Este hecho, además de las temperaturas muy altas, especialmente sobre la faz iluminada, descarta la existencia de agua apreciable y moléculas orgánicas. En lo que respecta a los exobiólogos, Mercurio no es un objeto de estudio de alta prioridad.

Venus, el segundo planeta desde el Sol, es de mayor interés. Es muy similar a la Tierra en tamaño y densidad, pero, a causa de su mayor proximidad al Sol, es más caliente que nuestro planeta. Los novelistas y algunos astrónomos alguna vez describieron a Venus como un paraíso tropical de junglas y pantanos, y rebosante de extrañas criaturas semejantes a reptiles y anfibios.

Los radiotelescopios y sondas interplanetarias han destruido este cuadro bastante atractivo de Venus. Desde el punto de vista exobiológico, Venus es realmente un lugar siniestro. La atmósfera es muy distinta a la de la Tierra; consiste principalmente en gas bióxido de carbono, con pequeñas cantidades de agua y de monóxido de carbono. Nubes densas, cuya composición exacta aún desafía el análisis, impiden ver la superficie del planeta.

Por ello no conocemos mucho acerca de la superficie de Venus. Sabemos, sin embargo, que las temperaturas y presiones del aire allí son intolerables para los seres humanos. Las temperaturas fluctúan entre 480° y 540°C, calor suficiente como para fundir plomo. Las presiones atmosféricas son de 60 a 100 veces mayores que aquellas de la superficie de la Tierra. El extremado calor hace que sea imposible la existencia de agua y compuestos orgánicos en la superficie: por tanto, no es habitat apto para la vida. Quizá las nubes sean albergue para organismos; se cree que hay una capa de nubes con temperaturas de sólo 15° a 32°C, que contienen algo de agua y oxígeno. Si existe vida allí, requeriría una ecología completamente surgida del aire, sin equivalencia en la Tierra.

¿VIDA EN MARTE?

Marte es el cuarto planeta desde el Sol. Su diámetro es sólo la mitad del de la Tierra. Pero tiene muchas características parecidas a las terrestres, y éstas llevaron a los astrónomos a creer que podría albergar vida. Inclusive se pensó que en Marte había seres inteligentes.

Estas ideas debieron cambiarse después que sondas no tripuladas volaron más allá de Marte y enviaron información a la Tierra, revelándolo como un mundo seco, frío, más parecido a la Luna que al nuestro, con amplias zonas cribadas de cráteres. La densidad de la atmósfera de Marte es increíblemente baja, como la atmósfera de la Tierra a una altura de aproximadamente 50 km sobre el nivel del mar. La atmósfera consta principalmente de bióxido de carbono, con un vestigio de agua; no se han hallado ni oxígeno ni nitrógeno, pese a que probablemente existen, al menos en cantidades registrables. A causa de la tenuidad del aire, la radiación ultravioleta del espacio, nociva para la mayoría de los seres vivientes, alcanza fácilmente la superficie del planeta.

El agua puede haber cubierto alguna vez grandes partes de la superficie de Marte; actualmente se sabe que sus casquetes polares están compuestos de agua congelada, y muestras del suelo tomadas por las sondas Viking produjeron agua al ser calentadas. La temperatura superficial promedio de Marte es al-

rededor de 22°C más baja que la de la Tierra. En el ecuador marciano, la temperatura puede alcanzar a 24°C durante el día, y por la noche puede descender a −73°C.

Estas condiciones hacen que Marte sea un habitat poco favorable para la vida tal como la conocemos, pero los exobiólogos piensan que algunas formas vivientes pueden existir allí. Varios factores sostienen esta posibilidad. Por ejemplo, la radiación ultravioleta es fácilmente bloqueada por barreras físicas. Oxígeno, nitrógeno y agua pueden estar presentes en cantidades suficientes para sostener organismos resistentes, tales como bacterias, musgos y líquenes. La presencia de una capa de agua permanente congelada, o permafrost, en el suelo subsuperficial parece posible. Este hielo puede fundirse localmente de vez en cuando, proporcionando entonces agua para seres vivientes. El nitrógeno podría existir en pequeñas cantidades en la atmósfera; el suelo puede contener sales químicas con nitrógeno.

Los exobiólogos también basan las expectativas de vida marcianas en determinados experimentos de laboratorio. Han sometido bacterias, algas, musgos, líquenes, algunas plantas superiores e insectos, a medios marcianos artificiales. Un número sorprendente de organismos sobrevivió; ciertas bacterias y musgos también se comportaron satisfactoriamente bajo las duras condiciones, mientras había agua.

Otros experimentadores produjeron compuestos de carbono bajo condiciones marcianas también simuladas. Mezclas de monóxido de carbono y dióxido de carbono y vapor de agua en suelo estéril o vidrio pulverizado se expusieron a radiaciones ultravioleta. Resultaron varios compuestos de carbono más complejo, de los cuales se sabe que están relacionados con la formación de compuestos biológicos aún más complejos. Este experimento demuestra que las atmósferas de metano, amoníaco e hidrógeno no son siempre necesarias para la evolución química de la vida.

Pero los exobiólogos consideran tales experimentos con cautela. Es imposible reproducir las condiciones marcianas con precisión, porque sabemos muy poco acerca de ellas y nuestras técnicas son aún limitadas. También, a los organismos de ensayo en los experimentos exobiólicos se les dio suficiente agua y oxígeno para sobrevivir. No sabemos aún si realmente el oxígeno y el agua están presentes en cantidades suficientes en Marte para sustentar el tipo de vida terrestre. Sin embargo, nuevos datos pueden lograrse después del aterrizaje de las sondas Viking.

De todos modos, de estos experimentos se ha derivado una importante lección. Si la vida terrestre puede sobrevivir en Marte, nuestras sondas espaciales podrían fácilmente contaminar aquel planeta. Para evitarlo, deberíamos esterilizar completamente cualquier nave que aterrice en Marte.

LOS PLANETAS EXTERIORES Y LA VIDA

Durante muchos años, Júpiter, Saturno, Urano y Neptuno no han sido considerados aptos para la vida. Si examinamos los hechos acerca de estos planetas, esta actitud es razonable.

Por sus grandes distancias del Sol, los planetas exteriores son extremadamente fríos. Los astrónomos no están seguros de que todos tengan cuerpos sólidos. A pesar de sus inmensas dimensiones y masas, Júpiter, Saturno, Urano y Neptuno son menos densos que la Tierra. De existir, el núcleo sólido de un planeta exterior sería pequeño, comparado con su volumen total. Posiblemente consistiría en gases líquidos y congelados, metales y quizá roca.

Se puede advertir por qué los científicos hallaron difícil representarse la existencia de vida en los planetas exteriores; pero hoy, por varias razones, los exobiólogos están cambiando sus puntos de vista. Primero, ahora saben que las atmósferas de amoniaco, metano e hidrógeno pueden ser viveros de sustancias bioquímicas y, por tanto, de gran importancia para la evolución química. Más aún, los planetas exteriores no siempre son tan fríos como alguna vez pensaron los astrónomos. Estudios instrumentales de estos mundos muestran que algunas de sus capas de nubes pueden tener una temperatura relativamente moderada.

La atención se ha centrado particularmente en Júpiter. Acaso tenga una fuente interna de energía calórica que lo haga más cálido de lo que podría ser si sólo lo calentara el Sol. Su manto de nubes multicolores posiblemente contenga gran variedad de compuestos orgánicos. Los experimentos de laboratorio han mostrado que actualmente pueden muy bien estar produciéndose compuestos orgánicos y biológicos en la atmósfera de Júpiter.

Plutón, el planeta más exterior de nuestro Sistema Solar, es de escaso interés exobiológico. Se sabe muy poco acerca de él. Parece ser un pequeño cuerpo sólido, intensamente frío a causa de su inmensa distancia del Sol. Cualquier atmósfera que pudiera tener, estaría permanentemente congelada.

Otros cuerpos del Sistema Solar pueden ser de importancia exobiológica. Ya hemos tratado los meteoritos. Los asteroides, probablemente una fuente importante de meteoritos, podrían contener sustancias químicas orgánicas. Los cometas también tienen gran interés, y estudios espectroscópicos ya han revelado que en éstos hay moléculas que contienen carbono.

EXPLORACIÓN EXOBIOLÓGICA

El conocimiento exobiológico de los planetas, y posiblemente de otros cuerpos de nuestro Sistema Solar, aumentará a medida que más naves espaciales lleven instrumentos científicos y hombres al espacio. Las sondas espaciales no tripuladas ya han sobrevolado y circundado a Marte y a Venus, y descendido en ellos. Estas naves no fueron diseñadas para detectar la presencia de seres vivientes en estos dos planetas. Tampoco han hallado un rastro inequívoco, pasado o presente, de vida en la Luna.

La futura nave espacial que descienda en Venus, y especialmente en Marte, llevará instrumentos que habrán de poder identificar sustancias químicas biológicas y posibles formas de vida. Otra nave ha sido enviada a Mercurio y también a los planetas exteriores: Júpiter, Saturno, Urano, Neptuno y Plutón. En 1975, la NASA lanzó un par de naves espaciales Viking, que aterrizaron en Marte en 1976. Una parte de cada Viking circundó a Marte y lo estudió con instrumentos de gran alcance. La otra parte, la de descenso, se posó suavemente en la superficie del planeta. Todas han enviado informaciones acerca de Marte a la Tierra.

A bordo de los Viking que descienden se realizan experimentos y hay instrumentos para detectar moléculas biológicas y microorganismos, que se suponen basados en compuestos de carbono. Los sensibles analizadores de cada parte de descenso descubrirían la presencia de microorganismos en el suelo marciano por las actividades químicas y los cambios que éstos producen.

La parte que descendió en Marte llevó a cabo automáticamente cuatro experimentos principales:

(1) Se expuso una muestra del suelo marciano a elementos químicos radiactivos de la Tierra. Si se descubre que el suelo contiene organismos, y si estos organismos emplean elementos químicos durante sus procesos vitales, esta actividad será seguida por los instrumentos detectores de radiación.

(2) Se probó la atmósfera sobre una muestra de suelo marciano. Si se hallan cambios químicos mensurables en la atmósfera, dichos cambios estarán causados probablemente por organismos.

(3) Se analizó un poco de suelo para ver si una muestra suministrada de carbono radiactivo se incorporaba fotosintéticamente a organismos marcianos.

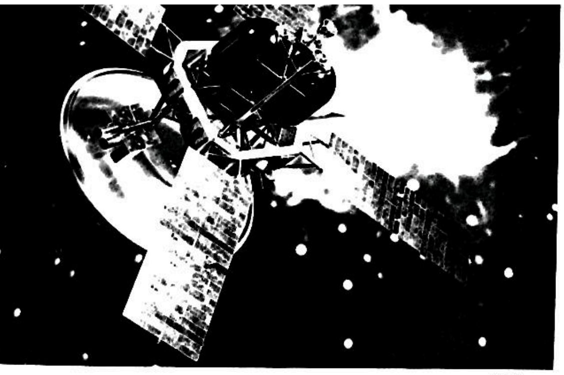
(4) Se analizó químicamente una muestra de suelo en busca de la presencia de moléculas orgánicas y de agua.

El análisis preliminar de los datos arrojó algunos resultados sorprendentes, que pueden indicar cierto grado de actividad biológica en el suelo.

CIVILIZACIONES EN EL ESPACIO EXTERIOR

Sería realmente un triunfo encontrar cualquier clase de vida en otro planeta. Más interesante aún sería descubrir vida inteligente. ¿Cómo estableceríamos contacto con seres inteligentes en otros mundos? La idea de comunicarse con criaturas de otros planetas surgió cuando el hombre advirtió por primera vez que los planetas eran mundos distintos. Los primeros astrónomos que usaron telescopios poseían muy poco conocimiento de las condiciones en los diferentes planetas del Sistema Solar. Por tanto, imaginaron que éstos estaban habitados por seres inteligentes, y dedujeron que esas criaturas querrían comunicarse con nosotros si descubrían que la Tierra también albergaba vida inteligente.

De vez en cuando se han propuesto esquemas de comunicación interplanetaria. Todos ellos incluían el uso de enormes letras, símbolos, números o figuras geométricas que se suponía podrían verse desde los mundos más cercanos a la Tierra: Marte, Venus y la Luna. Por ejemplo, un hombre sugirió cavar largas zanjas en la Tierra en forma de diferentes signos, que se llenarían con petróleo y se en-



NASA

Un modelo de la nave espacial Viking en vuelo simulado. Dos de estas sondas circundaron a Marte y aterrizaron en él, tomando muchas fotografías de la superficie del planeta, realizando experimentos atmosféricos y excavando y analizando muestras del suelo en procura de posibles indicios de vida.

cenderían. Probablemente observadores inteligentes de planetas vecinos advertirían el mensaje, y se enterarían que la Tierra está civilizada.

Tales ideas no eran factibles y escasas personas las tomaron seriamente. Más aún, el esfuerzo se habría perdido, porque ahora sabemos que la vida inteligente quizá no existe en nuestro Sistema Solar, salvo en la Tierra.

Con el advenimiento de la radio a fines del siglo XIX, la idea de enviar mensajes electrónicos o señales a otros planetas fue ganando terreno. En la década del veinte se emitieron señales al espacio. Ocasionalmente, un radiooperador anunció que había recibido señales que no podían haber emitido estaciones terrestres. Pero la mayoría de estos "mensajes" provenía quizá de un equipo defectuoso o de electricidad atmosférica. Otros podrían haber sido radioondas del espacio exterior. La existencia de radioondas cósmicas se probó en 1938, pero, por lo que sabemos, eran de origen natural. (Como mencionamos antes, los radioastrónomos regularmente sintonizan esas ondas cósmicas para estudiar la naturaleza de los cuerpos celestes que las emiten.)

En los últimos años de la década del 60 sorprendió al mundo la noticia de que se estaban recibiendo pulsos electromagnéticos de extraordinaria regularidad desde ciertos puntos del espacio exterior. Los periódicos y revistas publicaron teorías acerca de supercivilizaciones cósmicas que trasmitían mensajes, o de naves espaciales extrañas guiadas por faros también espaciales; pero los astrónomos pronto descubrieron que las señales son naturales; provenían de emisiones de estrellas poco comunes llamadas pulsares.

Dichos pulsos son simplemente un ejemplo de las muchas clases de radiación electromagnética en el espacio. La radiación va desde rayos gamma ultracortos hasta radioondas muy largas. La tarea de seleccionar entre todas estas longitudes de onda para encontrar un mensaje inteligible sería muy difícil. Algunos expertos sugieren sintonizar la longitud de onda emitida por moléculas de hidrógeno (21 cm), que es la más común de las radioondas cósmicas. Alguna civilización podría utilizar esta longitud de onda universal para enviar mensajes a través del espacio, pero otras longitudes de onda le servirían también.

Las radioondas y otras formas de radiación electromagnética, como la luz, se desplazan a través de un vacío a unos 300.000 km por segundo. A esta velocidad la radiación viajaría aproximadamente 9,6 billones de kilómetros en un año. Esta distancia se denomina un año luz. Para nosotros, en la Tierra, parece una distancia imposible. ¡Nada puede estar tan lejos! Sin embargo, la estrella más próxima a nuestro Sistema Solar, Alfa del Centauro, está a 4,4 años luz de nosotros. Por tanto, si enviáramos un radiomensaje a un planeta que gire en torno de Alfa del Centauro tardaría alrededor de 4,4 años en llegar a destino. La respuesta echaría al menos el mismo tiempo para llegar a nosotros. Luego una comunicación en ambos sentidos entre la Tierra y un planeta de Alfa del Centauro ¡llevaría cerca de nueve años para un solo intercambio de mensajes!

Muchas estrellas y sus planetas están a millones o miles de millones de años luz de la Tierra. Por tanto, la comunicación en dos sentidos con seres inteligentes de esos mundos es imposible.

Cuando nos llegue un mensaje de una de esas civilizaciones distantes, éstas podrían haber dejado de existir hace tiempo. Sin embargo, recibir simplemente un mensaje inteligible o señal del espacio exterior sería emocionante. Probaría que existen seres inteligentes, o existieron, en algún lugar del Universo.

A pesar de las dificultades implicadas, los científicos han intentado escuchar señales inteligentes del espacio exterior. En 1960-61, el Dr. Frank Drake, en el Observatorio Nacional Radioastronómico de Green Bank, Virginia Occidental, usó un radiotelescopio para este propósito. Orientó la antena de 26 metros del telescopio hacia dos estrellas próximas que emiten radioondas. El experimento se denominó Proyecto Ozma. El Dr. Drake no logró descubrir ningún esquema inteligente en el radiorruido que recibió desde estas estrellas, y el Proyecto Ozma finalizó en unos pocos meses.

Experimentos similares en la Unión Soviética tuvieron resultados no confirmados. Durante un tiempo, un investigador detectó señales, con alguna regularidad, provenientes de una parte del cielo. Pero este hecho no lo confirmaron otros radioastrónomos.

¿ESTAMOS SOLOS?

Los científicos, sin embargo, no han perdido las esperanzas. El Centro de Investigaciones Ames de la NASA y la Universidad de Stanford emprendieron un estudio denominado Proyecto Cíclope. Su propósito era determinar las posibilidades de recibir mensajes del espacio exterior. El Proyecto Cíclope sugería establecer una formación amplia de radiotelescopios con hasta 10.000 antenas cóncavas, cada una de 30 metros de diámetro, dispuestas en un área de 16 a 32 km de ancho.

Dicha formación podría detectar dispersiones de microondas y radioondas de civilizaciones planetarias distantes hasta 100 años luz de la Tierra. La dispersión ocurre cuando algunas de las ondas usadas para radio, televisión y radar escapan de un planeta al espacio exterior. Esta dispersión puede viajar muchos años luz. Ondas similares escapan diariamente de nuestra atmósfera; es posible que alguna civilización distante en este preciso momento esté sintonizando la dispersión electrónica de la Tierra.

Queda por saber si existe vida, inteligente o no, en otros planetas. A muchos les disgusta la idea de que haya seres vivientes en cualquier otra parte del Universo. Otros reciben con agrado esta posibilidad.

NASA

Foto de Júpiter tomada durante un vuelo de reconocimiento del Pioneer sobre el planeta. Los exobiólogos piensan que la atmósfera de Júpiter, con metano, amoníaco e hidrógeno, podría quizás estar produciendo elementos quimicobiológicos.

OBJETOS VOLADORES NO IDENTIFICADOS

El 24 de junio de 1947, un comerciante de Boise (Idaho), Kenneth Arnold, piloteaba un avión privado cerca del monte Rainier, Washington. Súbitamente lo sobrecogió ver un grupo de naves de extraño aspecto que realizaban asombrosas maniobras "Volaban muy cerca de las cimas de las montañas", dijo después, "volaban... como si estuvieran enlazadas unas a otras. Las observé por unos tres minutos: formaban una cadena que tenía por lo menos siete kilómetros de largo, y evolucionaban sobre las altas cumbres. Estos objetos eran chatos como sartenes y tan brillantes que reflejaban el sol a manera de espejos. Nunca vi nada tan veloz."

Cuando Arnold llegó a su destino –Yakima, Washington– y contó lo que había visto, causó sensación. Los "platillos voladores" que Arnold había visto se convirtieron en tema apasionante de conversación. A los pocos días, las naves volantes semejantes a discos habían sido vistas por observadores en otras partes del país y pronto se extendió el temor a los platillos voladores.

Los escurridizos discos, de los que entonces se hablaba tan frecuentemente, eran simplemente los últimos de una serie de objetos misteriosos del firmamento que han asombrado a los hombres desde tiempo inmemorial. La mayoría de las apariciones anteriores había sido aclarada sin lugar a dudas: eran meteoros, cometas, fenómenos atmosféricos, etc.; unas pocas nunca se explicaron satisfactoriamente. Algunas despertaron, en un momento dado, amplio interés durante semanas o meses: luego se olvidaron.

La epidemia de platillos voladores iniciada por el relato de Arnold alcanzó proporciones bastante respetables en 1947; pero hacia fines del año el interés público parecía haber declinado. Lo revivió poderosamente una tragedia ocurrida el 7 de enero de 1948. A primera hora de la tarde de aquel día, observadores de la Base Godman de la Fuerza Aérea, en Kentucky, vieron un objeto misterioso que volaba sobre sus cabezas. Parecía "un cono de helado coronado de rojo". Se pidió a cuatro pilotos de la Guardia Nacional, en aviones F-51, que investigaran la extraña máquina voladora. El capitán Thomas F. Mantell, jefe del vuelo, comunicó por radio a la torre de control que "estaba acercándose para lograr una buena visión". Al cabo de un rato informó que la cosa parecía metálica y de tamaño enorme. "Ahora está subiendo y adelantándose tan rápidamente como yo... Es decir, a 54[illegible] kilómetros por hora. Voy a elevarme hasta 6.000 metros y si no me aproximo más abandonaré la caza."

Después de este último informe, recibido a las 3,15 p.m., no hubo otro contacto radial con Mantell. Ese mismo día, su cuerpo fue encontrado entre los restos de su avión cerca de Fort Knox. La explicación oficial de la Fuerza Aérea norteamericana fue que Mantell se había desmayado a 6.000 metros por falta de oxígeno y había muerto por asfixia antes de estrellarse. El objeto perseguido por Mantell fue primero identificado como el planeta Venus, pero indagaciones subsiguientes demostraron que éste no se hallaba aquel día en la zona del cielo donde se observó el misterioso objeto.

MUCHOS RELATOS DE VISIONES

Se produjo entonces una nueva ola de visiones de platillos voladores. Algunas de las naves misteriosas fueron avistadas por aviones; otras, por observadores en tierra; una porción de ellas se captó en pantallas de radar. Entre los informantes se incluían varios profesionales expertos: localizadores de proyectiles dirigidos, operadores de radar, pilotos aéreos-comerciales, pilotos de la Fuerza Aérea norteamericana, reguladores del tránsito aéreo, observadores meteorológicos. Los platillos se vieron en muchas partes distintas de los Estados Unidos, pero particularmente en las desiertas zonas del sudoeste. También hubo relatos de apariciones en Canadá, México, América del Sur, Europa, el Lejano Oriente, Australia, África, Hawai, Groenlandia y la Antártida.

La Fuerza Aérea de los Estados Unidos, decidida a descubrir el misterio, inició una serie de investigaciones. Los investigadores dieron a las enigmáticas naves el nombre oficial de "objetos voladores no identificados": OVNI.

Los OVNI descritos por los observadores fueron de muy distintos tipos. Predominaban los platillos voladores o discos. Había discos grandes de hasta 30 metros de diámetro, discos de tamaño mediano y minúsculos discos de sólo algunos centímetros. Estas naves realizaban maniobras asombrosas. A veces permanecían quietas en el aire, luego se lanzaban hacia el cielo, daban vueltas bruscas y cambiaban su curso con increíble rapidez. Alcanzaban las más fantásticas velocidades, hasta miles de kilómetros por hora en algunos casos.

Había asimismo naves en forma de cohetes que medían entre 30 y 300 metros de largo, también capaces de tremendas velocidades. En algunos casos, estos cohetes parecían servir como naves madres para los discos. Además de los discos y los cohetes había brillantes bolas de fuego verdes que se movían silenciosamente por los cielos tan velozmente como meteoros.

SE VEN SERES EXTRAÑOS

Ciertos observadores aseguraron haber visto tripulantes de estas naves espaciales. En un libro de 1950 titulado *Detrás de los platillos voladores*, Frank Scully habló de ciertos hombres de Venus cuyos cadáveres se habían hallado después que su aparato cayó a tierra en Nuevo México y Colorado. Según el informante de Scully, "un señor Newton", los hombres de Venus eran "seres minúsculos, de 97 a 112 centímetros de alto. Vestían trajes de 1890, de una tela que no era lana ni algodón, pero que no podía rasgarse". Según Newton, la Fuerza Aérea se había llevado los cuerpos.

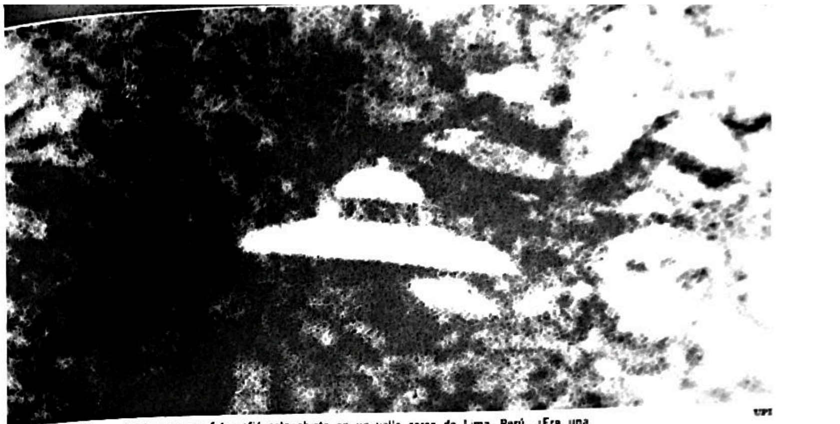
Aparentemente, no todos los tripulantes eran enanitos. El 12 de setiembre de 1952, una mujer, tres niños y un joven guardia nacional divisaron un platillo volador cerca de Sutton, Virginia Oeste. Poco después dieron con un repulsivo gigante, al parecer miembro de la tripulación. Medía 2,70 metros de altura y tenía rostro rojo y ojos saltones, separados cerca de 30 centímetros. Cuando este monstruo se dirigió hacia los espantados observadores con un sonido sibilante, ellos huyeron.

Hubo muchísimas fotografías de platillos. Algunas de éstas eran evidentemente fraguadas. En otros casos se habían producido en negativos imágenes raras por fallas del equipo o la técnica. Sin embargo, otras pocas no pudieron explicarse tan fácilmente.

Después de 1953, las apariciones de platillos voladores fueron menguando. Quizá sería más exacto decir que en periódicos y revistas aparecían menos relatos de objetos voladores no identificados, acaso porque los editores pensaron que tales relatos ya no constituían noticias novedosas. En el período 1965-67 hubo un marcado aumento de información sobre apariciones, y luego una nueva disminución.

EN BUSCA DE EXPLICACIONES

¿Cómo podemos explicarnos estos misteriosos objetos no identificados del cielo? ¿Son



UPI

Un arquitecto peruano fotografió este objeto en un valle cerca de Lima, Perú. ¿Era una nave del espacio exterior? ¿Un truco? ¿O lo causaban fenómenos normales?

United Press

Concepción de un platillo volador realizada por un artista, basada en descripciones hechas en primitivos relatos de apariciones.

realmente naves voladoras mucho más avanzadas que cualquier otra conocida antes por el hombre? Después de todo –se ha argüido– el hombre ha viajado a la Luna. ¿No será posible que habitantes de otros sistemas estelares, poseedores de una tecnología superior, hayan perfeccionado naves que puedan conducirlos a nuestro Sistema Solar?

Poco después que Kenneth Arnold vio aquellos "platillos voladores" en 1947, la Fuerza Aérea norteamericana comenzó a rastrear todas las informaciones relativas a los OVNI en lo que vino a conocerse como Proyecto Libro Azul. En 1966 contrató a un grupo de 36 científicos, bajo la dirección del físico Edward U. Condon, de la Universidad de Colorado, para conducir una investigación independiente de tales apariciones. El estudio culminó con la publicación de un voluminoso informe que determinaba que no tendría razón de ser una nueva investigación sobre los OVNI. En 1969, la Fuerza Aérea dio por terminado el Proyecto Libro Azul, aunque unos 1.000 informes de sus archivos quedaron sin explicar.

Las referencias a las apariciones de OVNI decayeron bruscamente, pero la controversia sobre su causa no terminó. Algunos científicos continuaron pidiendo una investigación exhaustiva de la Fuerza Aérea. El Dr. James McDonald, por ejemplo, señaló que los autores del Informe Condon "admitían que casi un tercio de los 90 casos que habían examinado no se podía explicar. No está claro cómo esto justificaba la conclusión de que no se necesitaba un nuevo estudio".

Luego, en 1973, se produjo un súbito resurgimiento de las apariciones de OVNI. Uno de los relatos más interesantes fue el de dos hombres de Missisipi, quienes contaron haber sido embarcados en una nave espacial azul por seres que tenían piel rugosa plateada y manos como pinzas de cangrejo. El astrónomo J. Allen Hynek, que fuera consultor en el Proyecto Libro Azul, examinó a los hombres bajo hipnosis y tuvo la impresión de que estaban "diciendo la verdad, fuera de una duda razonable".

Otras apariciones tuvieron pronto explicaciones convencionales. Por ejemplo, una serie de OVNI eran realmente globos atmosféricos plateados que reflejaban la luz solar a sus 26 kilómetros de altura. Otros resultaron ser fenómenos meteorológicos tales como nubes bajas o niebla. No obstante, una encuesta de norteamericanos adultos demostró que el 51 por ciento creía que los OVNI eran objetos reales, no simples imaginaciones de la gente. Y un 11 por ciento dijo que había visto OVNI.

A fines de 1973, El Dr. Hynek, que es ahora presidente del Departamento de Astronomía en la Universidad del Noroeste, anunció la formación de un Centro para el Estudio de los OVNI. "Durante un cuarto de siglo el fenómeno OVNI ha sido tema de malas informaciones, errores, ridículo, burla y enfoques no científicos", dijo. "El hecho de persistir referencias desde muchos pasíes, presenta un misterio que exige explicación."

ALGUNAS EXPLICACIONES CONVENCIONALES

Según se ha indicado, muchos platillos voladores no son realmente naves espaciales. Numerosos científicos, como el astrónomo Donald H. Menzel, de Harvard, sostienen que muchas de las apariciones misteriosas han sido causadas por fenómenos meteorológicos. En algunos casos, las apariciones semejantes a platillos pueden deberse a la reflexión de la luz desde formaciones de cristales de hielo en la atmósfera. Si esos cristales están cayendo o suspendidos en el aire, la reflexión de la luz del Sol o de la Luna puede producir efectos sorprendentes. Por ejemplo, provocar la aparición de un par de halos concéntricos alrededor del Sol. A veces, parte de los halos llega a ser casi tan brillante como el propio Sol y a formar falsos soles, llamados también parhelios. Por la noche se desarrollan halos alrededor de la Luna, formando falsas lunas o paraselenes, que podrían bien asemejarse a extrañas naves espaciales.

En otros casos, las apariciones de platillos voladores quizá se hayan debido a que el aire puede actuar como una lente deformante, produciendo los efectos conocidos como espejismos. Los rayos de luz se reflejan o encorvan de varios modos al pasar por capas de aire de distinta densidad. El efecto es más notable cuando capas de densidad abruptamente contrastante se hallan en estrecho contacto.

Si una capa de aire caliente está cerca de la Tierra, la imagen del cielo puede proyectarse contra ésta, dando a veces la ilusión de charcos o lagos a la distancia. Si cerca de la Tierra hay una capa fría (como en el desierto por la noche) la imagen de la Tierra puede proyectarse contra el cielo. Las luces distantes, tales como las de los automóviles o de una ciudad, parecerán flotar en el aire. Si el aire es turbulento, esas luces aparentemente danzarán de acá para allá. Las condiciones que causan espejismos ópticos pueden producir también espejismos de radar.

En algunos casos, la gente puede haber avistado aeronaves, pero no los exóticos objetos que ha imaginado. Quizás haya visto aviones comunes volando a tal altura que sólo era visible el reflejo del Sol en el fuselaje.

Otras personas han confundido cometas (barriletes o papalotes) o globos meteorológicos con OVNI. En 1951, el físico Urner Liddel afirmó que la mayoría de los OVNI mencionados anteriormente eran en realidad globos sonda, grandes naves plásticas no tripuladas que se usan para elevar instrumentos meteorológicos. Estos globos alcanzan grandes alturas y frecuentemente vuelan grandes distancias; algunos han cruzado el Atlántico. Cuando se ven contra el fondo del cielo a menudo parecen exactamente discos. Algunas autoridades creen que el capitán Mantell se estrelló al perseguir un globo sonda. Fue en aquel momento un asunto secreto, y al público no le llegó información al respecto.

Cuando esta información se hizo pública pocos años después, resolvió un sobrecogedor misterio. El 17 de marzo de 1950, los habitantes de Farmington, Nuevo México, se espantaron hasta casi perder el juicio cuando miles de "platillos voladores" se cernieron sobre la ciudad durante casi una hora. Sólo tiempo después se reveló que un globo sonda había explotado sobre Farmington a 18 kilómetros de altura, dispersando en el aire miles de pedazos de plástico, que invadieron la zona.

¿PRODUCTOS DE LA IMAGINACIÓN?

Científicos que estudian el comportamiento se ocupan también de los informes relativos a los OVNI. El sociólogo Robert Hall señala que "el cielo, especialmente el nocturno, está lleno de estímulos equívocos y la gente suele sentir una urgente necesidad de reducir la ambigüedad . . . mediante explicaciones basadas en algo conocido". Así, algunas personas, influidas por el general "sistema de fe" que se ha desarrollado alrededor del fenómeno OVNI, trasforma estímulos naturalmente explicables en una nave del espacio exterior.

El Dr. Lester Grinspoon sugiere que las visiones de OVNI acaso estén ligadas con las tensiones de la vida moderna, con los "cada vez más angustiosos tiempos en que vivimos". Bajo tales tensiones –dice– es común que la gente tenga ilusiones o desilusiones, una fantasía sustitutiva "que proporcione lo que la realidad ha negado".

El Dr. Carl Sagan, de la Universidad de Cornell, ha propuesto que "ciertas necesidades psicológicas se satisfacen mediante la creencia en seres superiores de otros mundos". Y hoy –dice– los visitantes extraterrestres son una idea que está de moda. El Dr. Sagan, no obstante, previene contra el rechazo de la hipótesis extraterrestre. No existe suficiente evidencia –a su entender– para excluir la posibilidad de que algunos OVNI sean naves espaciales de civilizaciones avanzadas que viven en algún otro lugar del Universo.

En 1794, la Academia de Ciencias francesa se negó a aceptar la posibilidad de que los meteoritos se originaran fuera de la atmósfera terrestre. "Los meteoritos no encajaban en la creencia establecida; por tanto, los científicos negaron su existencia", dijo el Dr. Hall. ¿Es esto mismo cierto en cuanto a los platillos voladores del espacio exterior?

COMPUTADORAS Y MATEMÁTICAS

Izquierda: Sperry Univac, división de Sperry Rand Corporation
Arriba: General Automation Inc.

***Las computadoras** han llegado a ser parte **integral** de la vida moderna. Izquierda: una **computadora** grande con unidad central de **procesamiento** (al fondo); discos magnéticos **de la unidad** de almacenamiento (arriba **izquierda**) y salida o unidad impresora al **frente**). Arriba: un chip de procesamiento de **datos** de una moderna microcomputadora.*

38-43	**Introducción a las matemáticas**
44-51	**Los números**
52-57	**Aritmética**
58-71	**Algebra**
72-85	**Geometría plana**
86-91	**Geometría del espacio**
92-96	**Trigonometría**
97-106	**Geometría analítica**
107-109	**Geometría no euclidiana**
110-119	**Estadística**
120-129	**Probabilidad**
130-136	**Teoría del juego**
137-143	**Cálculo**
144-151	**Teoría de los conjuntos**
152-158	**Números binarios**
159-167	**Procesamiento de datos**
168-180	**Computadoras**
181-189	**Programación de computadoras**
190-191	**Procesamiento de palabras**

INTRODUCCIÓN A LAS MATEMÁTICAS

La palabra "matemáticas", del griego *mathemata*, significa "cosas que se aprenden". Puede parecer extraño aplicar esta frase a un campo particular del conocimiento, pero deberíamos señalar que, para los antiguos griegos, las matemáticas incluían, además del estudio de los números y el espacio, astronomía y música. En la actualidad, naturalmente, no pensamos que la astronomía y la música sean temas matemáticos; sin embargo, el alcance de las matemáticas es actualmente más extenso que nunca.

Las matemáticas modernas son un amplio campo de conocimiento con muchas subdivisiones. Existen, antes que nada, las matemáticas de los números, o cantidad. El ramo de la ***aritmética*** trata de números particulares como 3 ó 10½ ó 12,5. Cuando sumamos, restamos, multiplicamos o dividimos dichos números o sacamos sus raíces cuadradas, o cuadrados, efectuamos operaciones aritméticas. A veces no deseamos considerar números particulares, sino relaciones que podrán aplicarse a grupos completos de números. Estudiamos dichas relaciones en álgebra, otro ramo de la ciencia de la cantidad. En ***álgebra***, un símbolo, tal como la letra *a* o *b*, representa una clase completa de números. Por ejemplo, en la fórmula:

$$(a+2)^2 = a^2 + 4a + 4$$

la *a* representa cualquier número. La relación expresada por la fórmula se mantendría, tanto si *a* valiera 1 ó 5 ó 10 u otro número.

Las matemáticas también estudian las formas en el espacio, que deben imaginarse como un mundo de puntos, superficies y sólidos. Estudiamos las propiedades de las diferentes formas y las relaciones entre ellas, y aprendemos a medirlas. Esta ciencia del espacio se denomina ***geometría***.

La ***geometría plana*** trata de los puntos, líneas y figuras que están en un solo plano, una superficie con sólo dos dimensiones (Figura 1). El estudio del mundo tridimensional se denomina ***geometría del espacio*** (Figura 2). La trigonometría ("medición de triángulos") es un ramo de la geometría. Se basa en el hecho de que si se conocen ciertas partes de los triángulos, se pueden determinar las restantes y resolver muchos problemas distintos (Figura 3).

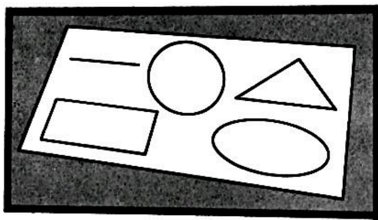
La ***geometría analítica*** combina álgebra y geometría, números generalizados y relaciones espaciales. Sitúa las figuras geométricas en el espacio. Explica circunferencias, elipses y otras figuras en fórmulas algebraicas. En la figura 4, por ejemplo, las líneas *XX* y *YY* forman ángulo recto y se cortan en el punto *O*. Para indicar la posición de la recta *AB* respecto de *XX* y *YY* usamos la fórmula $x = 2y$. Para dar la posición de la circunferencia en el diagrama, usamos la fórmula $x^2 + y^2 = 25$.

El ramo de las matemáticas denominado ***análisis matemático*** estudia las funciones. Si el valor de una determinada cantidad depende del que le asignamos a otra cantidad, decimos que la primera es una función de la segunda. Por ejemplo, sabemos que la circunferencia es siempre 3,14159 veces el diámetro. El valor 3,14159 se indica con la letra griega π, o pi. Frecuentemente, usamos la fórmula $C = \pi D$ (o sea $\pi \times D$), donde C es la circunferencia y D es el diámetro. En este caso, C, la circunferencia, es una función de D, el diámetro, pues su valor dependerá del valor que le asignemos a D. El ***cálculo integral*** trata el límite de los diferentes valores de una función variable. El ***cálculo diferencial*** determina la razón de variación de una función variable.

La ***estadística***, otro ramo de las matemáticas, incluye la acumulación y tabulación de datos expresados en cantidades, y las leyes generales basadas en dichos datos. La teoría de las probabilidades es parte importante de la estadística. Permite calcular qué pasa si las posibilidades de que ocurra un evento son iguales y si no lo son. La teoría se aplica a juegos de azar, tales como dados y póquer. Puede usarse, también, para predecir el número relativo de varones y mujeres que nacerán en un determinado momento y lugar.

Éstas son sólo unas pocas de las muchas subdivisiones de las matemáticas, las cuales, además de ser uno de los campos más extensos del conocimiento, representan un enfoque lógico aplicable a muchos otros campos distintos. Definen con exactitud lo que debe tratarse y establecen claramente los supuestos posibles. Entonces, sobre la base de las definiciones y las suposiciones, se forja una cadena de pruebas, siendo cada eslabón de la cadena tan fuerte como cualquiera de los demás. Los matemáticos hicieron gala de maravillosos poderes de imaginación al determinar lo que puede probarse y al elaborar ingeniosos métodos de prueba.

Parecería algo rebuscado pensar en las matemáticas como búsqueda de belleza. Sin em-



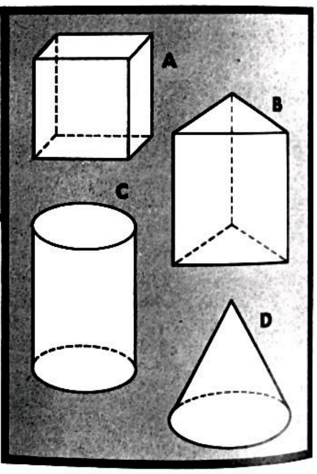


Figura 1. La línea recta, el círculo, el triángulo, el rectángulo y la elipse mostrados arriba se encuentran en un solo plano. Su estudio es parte de la geometría plana.

Figura 2. El cubo (A), el prisma (B), el cilindro (C), el cono (D) no están limitados por un solo plano. Estas figuras se estudian en geometría del espacio.

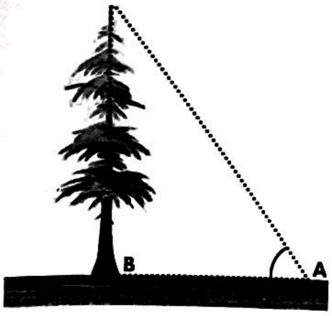


Figura 3. Si conocemos la distancia AB y el ángulo en A, podemos calcular la altura del árbol aplicando trigonometría, o medición de triángulos, un ramo de la geometría.

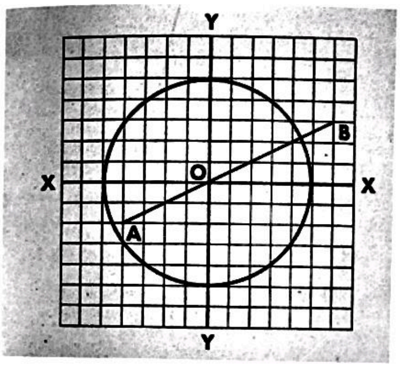


Figura 4. La línea XX está en ángulo recto con YY. La fórmula $x = 2y$ indica la posición de AB con respecto a XX y a YY. La fórmula $x^2 + y^2 = 25$ indica la posición de la circunferencia. Estamos tratando aquí de geometría analítica.

bargo, para muchos conocedores de este campo, los modelos matemáticos, combinados para formar un todo armónico, producirían efectos tan gratos como las combinaciones de colores de un pintor o las expresiones de un poeta. Como escribió Bertrand Russell, matemático importante del siglo XX, en sus *Principios de matemáticas:* "Las matemáticas, bien consideradas, poseen una belleza fría, austera, escultural, que no apela a ninguna parte débil de nuestra naturaleza, no posee lo decorativo de la pintura o la música, pero es sublimemente pura y capaz de la perfección severa del arte más superior."

También constituyen inagotable fuente de entretenimiento. Por generaciones, matemáticos y aficionados prepararon los llamados pasatiempos matemáticos, que van de problemas sencillos y construcciones, hasta acertijos para expertos, que a veces no pueden con ellos, pero que son un desafío encantador al ingenio; llevándonos a mundos de fantasía donde se "prueba" que $2 = 1$, o se construye un anillo mágico cuyo exterior es su interior (Figura 5).

LAS MATEMÁTICAS Y EL MUNDO EXTERIOR

Al desarrollar los diferentes ramos de las matemáticas, los iniciadores suelen deber mucho a la observación del ambiente. Se ha sugerido, por ejemplo, que los conceptos de "línea recta", "círculo", "esfera", "cilindro" y "ángulo" en geometría surgieron de objetos naturales: la "línea recta", del tronco de un árbol; el "círculo", del disco del Sol o la Luna; la "esfera", de algo redondo como un fruto, el "cilindro", del tronco de un árbol caído; el "ángulo", de los ángulos formados por los brazos o piernas en posiciones variables.

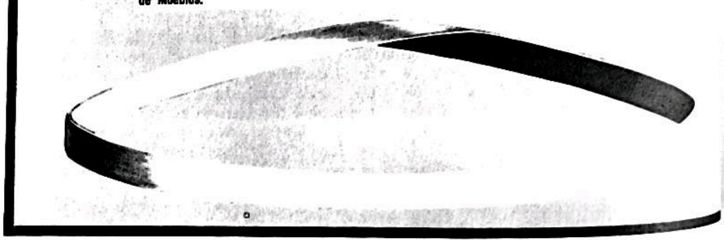
Los primeros matemáticos examinaron estas formas, relacionándolas. Primero, aplicaron los resultados a la solución de problemas prácticos, como la construcción de canales o la división de la tierra en lotes con fines de tasación. Luego, comenzaron a estudiar las relaciones entre varias formas geométricas para satisfacer su curiosidad intelectual y no para resolver problemas particulares. Con el tiempo, estructuraron una serie de conceptos puramente abstractos.

A medida que ciertos matemáticos desarrollan dichos conceptos en geometría, análisis matemáticos u otros ramos de las matemáticas se inclinan a volverle la espalda al mundo de la realidad. No vacilan en trabajar con ecuaciones de cuatro, o diez, o cualquier número de dimensiones. El matemático alemán G. F. B. Riemann construyó una geometría completa basada en el postulado, o supuesto, de que dos rectas nunca son paralelas. Otra geometría –que crearon independientemente N. I. Lobachevsky y J. Bolyai– partió del supuesto de que por un punto dado pueden trazarse al menos dos paralelas a una recta. El alemán Georg Cantor propuso una "teoría de los conjuntos" tan ingeniosa como inútil (al menos así pareció al principio) hasta que se le halló aplicación práctica.

Se ha indicado que el interés por la abstracción pura no carece de peligros. Como lo señaló el distinguido matemático Morris Kline en su libro *Las matemáticas y el mundo físico:* "Los matemáticos tienden a elevarse a las nubes del pensamiento abstracto, pero deben volver a la Tierra para alimentarse, o morir de inanición mental."

Sin embargo, aun sus especulaciones más abstractas pueden encontrar importantes aplicaciones; a veces después de muchos años. Así, la geometría riemanniana resultó inapreciable para Albert Einstein al desarrollar su teoría de la relatividad. La teoría de los conjuntos de Georg Cantor se aplicó en varios campos, incluyendo álgebra superior y estadística. Con respecto a las dimensiones múltiples, se han empleado, principalmente, en la inspección de tubos de televisión.

Figura 5. Este anillo mágico se realiza uniendo los dos extremos de una angosta tira de papel en la forma de un anillo, después de haber dado a un extremo una torsión de 180°. Si coloreamos un lado del anillo, descubriremos que se ha coloreado el anillo completo, por dentro y por fuera; en otras palabras, el interior es el exterior. El anillo se denomina cinta de Moebius.



APLICACIONES

No se requieren ejemplos extremos, como la geometría de Riemann y la teoría de los conjuntos, para demostrar cómo las matemáticas sirven a la humanidad. En efecto, están muy arraigadas en casi toda actividad humana, desde lo cotidiano hasta las avanzadas investigaciones de expertos en campos diferentes de la ciencia.

Todos somos un poco matemáticos. En nuestra vida usamos la aritmética a diario, al consultar un reloj para saber la hora; al calcular el costo de las compras y la vuelta que nos deben; al llevar los tantos en tenis, béisbol o fútbol.

Las operaciones de contabilidad de los negocios e industrias se basan en las matemáticas. El seguro es en gran medida una cuestión de interés compuesto y aplicación de la teoría de las probabilidades. Ciertos fabricantes usan el análisis matemático para utilizar mejor la materia prima. El piloto de un barco o de un avión usa la geometría cuando traza su carta de ruta. La mayor parte de la agrimensura se basa en la trigonometría. El ingeniero civil emplea aritmética, álgebra, análisis matemático y otros ramos de las matemáticas en su trabajo.

Las matemáticas también se aplican en las bellas artes, especialmente en arquitectura, pintura y música. Dan base a la perspectiva, sistema por el cual el artista representa en una superficie plana objetos y personas tales como aparecen en un mundo tridimensional. Podemos imaginar la perspectiva como compuesta por una serie de teoremas matemáticos. Uno, por ejemplo, establece que las rectas paralelas horizontales al alejarse en el plano del observador convergen en el denominado punto de fuga. La Figura 6 muestra la aplicación de este teorema al dibujar una habitación.

Las matemáticas desempeñan también un papel importante en música. El sistema de escalas y las teorías de armonía y contrapunto son básicamente matemáticas; también lo es el análisis de las calidades tonales de diferentes instrumentos. Ellas contribuyeron enormemente al diseño de pianos, órganos, violines y flautas y también aparatos reproductores, tales como fonógrafos o receptores de radio.

Las matemáticas son tan importantes en la ciencia y sirven en tantos de sus ramos que el notable matemático escocés-norteamericano Eric Temple Bell, la ha llamado la "reina y servidora de las ciencias". He aquí algunos ejemplos:

Las mediciones y otras técnicas matemáticas son vitales en el trabajo del físico. Éste utiliza el recurso matemático denominado representación gráfica para dar una imagen clara de la relación entre distintos valores; por ejemplo, entre la temperatura y la presión del vapor de agua saturado en la atmósfera (Figura 7). Las leyes de la física se enuncian a través de fórmulas algebraicas. Para dar la idea de que la velocidad de un cuerpo puede determinarse dividiendo la distancia recorrida por el tiempo empleado en recorrerla, el físico usa la fórmula $v = \frac{e}{t}$, donde v es la velocidad, e el espacio, o distancia recorrida, y t, el tiempo. El físico usa la geometría y la trigonometría en su análisis de las fuerzas y para establecer las leyes de la óptica, que es la ciencia de la luz.

Tanto el físico como el químico emplean continuamente operaciones aritméticas y algebraicas y representaciones gráficas. Este último también representa leyes con fórmulas algebraicas. Las reacciones se representan en forma de ecuaciones, que desde ciertos puntos

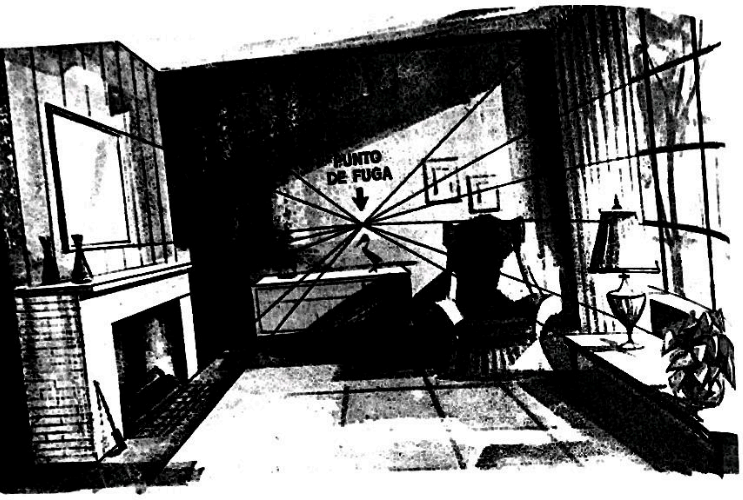
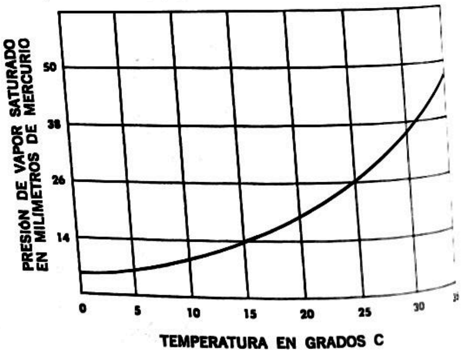


Figura 6. El efecto realista tridimensional en el grabado de arriba se obtiene haciendo converger las líneas horizontales que se alejan del observador en el punto de fuga.

de vista pueden considerarse como ecuaciones matemáticas; también usan logaritmos, una técnica matemática, para calcular el grado de acidez de una sustancia: el llamado valor pH. La geometría plana y del espacio se aplican en el estudio de la forma en que los átomos o los iones (átomos eléctricamente cargados) se combinan. Así puede mostrarse que los átomos de grafito forman una sucesión de hexágonos (figuras de seis lados) en una serie de planos paralelos dispuestos unos encima de otros (Figura 8); y los iones de cloro y sodio, que constituyen la sal común de mesa (cloruro de sodio), se colocan en las esquinas de una serie de cubos (Figura 9).

Figura 7. El gráfico de la derecha muestra cómo varía en la atmósfera la presión del vapor de agua saturado con la temperatura. Se alcanza el punto de saturación cuando la atmósfera no puede retener más vapor de agua a una temperatura dada.



Las matemáticas siempre estuvieron unidas a la astronomía. Desde los albores, los astrónomos midieron ángulos y arcos y realizaron muchísimos cálculos matemáticos al seguir el movimiento aparente del Sol, las estrellas, la Luna y los planetas. Hoy, ramos de las matemáticas, como aritmética, álgebra, geometría plana, geometría del espacio, trigonometría y análisis matemático, son tan útiles al astrónomo como el telescopio óptico, la cámara fotográfica, el radiotelescopio y otros dispositivos que usa en su trabajo.

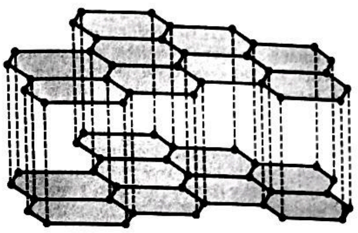


Figura 8. Cómo representa un químico la disposición de átomos del grafito, una forma de carbono. Los átomos forman una sucesión de hexágonos en una serie de planos paralelos dispuestos unos por encima de otros. Cada puntito negro representa un átomo de carbono.

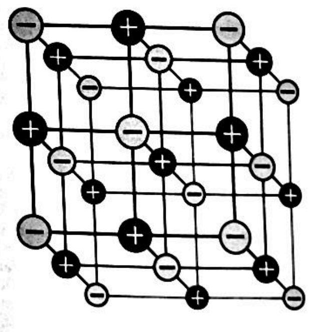


Figura 9. El químico representa los iones de sodio y de cloro que componen la sal de mesa como ubicados en las esquinas de una serie de cubos. Los círculos + representan los iones positivos de sodio; los círculos —, los iones negativos de cloro.

Parecería difícil aplicar fórmulas matemáticas al mundo infinitamente variado de lo viviente. Sin embargo, las matemáticas sirven aun en biología, la ciencia de la vida. Desempeñan un papel extremadamente importante, por ejemplo, en genética, que se relaciona con la herencia. Los genetistas usan la teoría de la probabilidad al calcular el porcentaje de individuos que tendrán, o no, rasgos parecidos en generaciones subsiguientes. Las matemáticas se han aplicado también a la comparación de formas relacionadas de vida. Por el método denominado análisis dimensional, los investigadores encontraron que la frecuencia del batir de alas de un pájaro puede sintetizarse en una fórmula. Es sorprendente ver que a menudo esta fórmula se aplica a pájaros muy diferentes. El análisis dimensional se usa también para estudiar los diagramas de crecimiento de ciertos animales y para determinar la proporción entre la duración de la vida de un determinado animal y el tiempo requerido por él para aspirar una vez.

Tal es, en resumen, el alcance de las matemáticas. En los artículos siguientes presentaremos algunos de los campos más importantes de esta ciencia, incluyendo aritmética, álgebra, geometría plana y del espacio, trigonometría, geometría analítica y análisis matemático, y diremos de qué tratan estos campos y a qué se aplican.

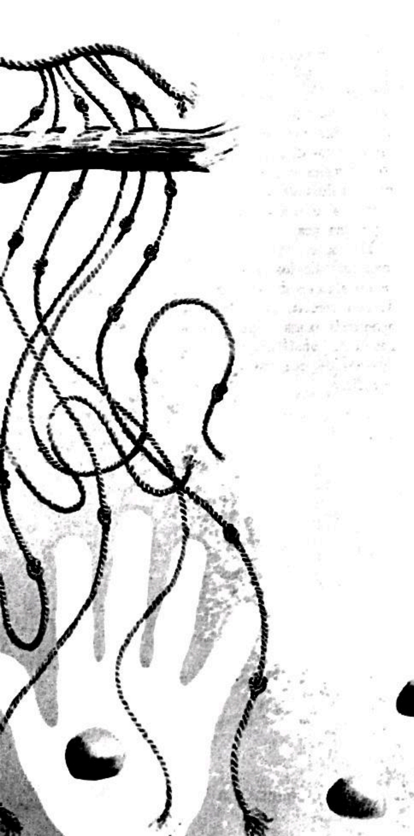
LOS NÚMEROS

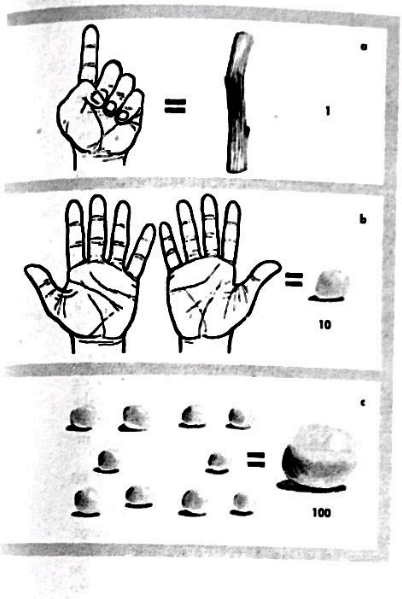
En el Museo Ashmoleano de Oxford, Inglaterra, hay una insignia real egipcia donde se registra la existencia de 120.000 prisioneros y un botín consistente en 400.000 bueyes y 1.422.000 cabras. Esta inscripción, anterior al año 3400 a.C., muestra que en la antigüedad la gente había aprendido a escribir cifras grandes. Por supuesto, los comienzos del uso de los números deben remontarse a mucho antes de los egipcios.

El hombre primitivo de las cavernas no tenía que saber gran cosa acerca de la numeración o cualquier otro tipo de matemáticas para sobrevivir. La cueva era un hogar que le quedaba a mano; el alimento se encontraba en los árboles o plantas, o podía cazarse con armas primitivas. Sin embargo, cuando se comenzó a reunir a los animales en rebaños, y especialmente cuando una familia entraba en relaciones sociales con otras, se volvió necesario decidir cuánto pertenecía a una persona y cuánto pertenecía a un vecino. Probablemente, al principio bastaba con usar conceptos como poco, o algo, o mucho. Luego, se hizo necesario tener medios más definidos para determinar este cuánto. La gente aprendió a contar, y éste fue el comienzo de las matemáticas.

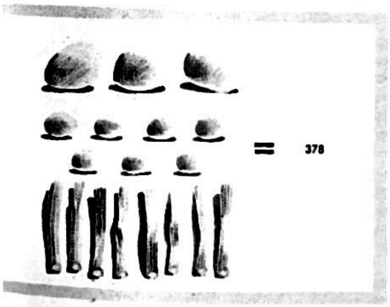
Al principio, una persona podía contar el número de animales de un rebaño colocando una piedrecilla en el suelo o haciendo un nudo en un cordel por cada animal. Cada piedrecilla en el creciente montón o cada nudo en el cordel representaría un único animal. Luego, un hombre quizás usó los diez dedos para sus cálculos. Podemos presumir que cuando se habían contado los diez dedos se colocaba una piedra pequeña a un lado para representar estos primeros diez; los dedos se usarían entonces para contar otras decenas; y se agregaría otra piedra a la primera; luego se emplearían de nuevo los dedos para contar otros diez, y así sucesivamente. Cuando las piedras en la pila igualaban el número de dedos contados, ellas representaban diez decenas. La pila de diez piedras se haría a un lado, y una piedra más grande se colocaría en su lugar para indicar diez decenas o una centena (Figura 1). Así, tres piedras grandes, siete piedras chicas y ocho palitos, en reemplazo de ocho dedos, representarían tres centenas, siete decenas y ocho unidades; en otras palabras, 378 (Figura 2).

Los diez dedos en este caso marcarían el número máximo en los cálculos de una persona lo que podríamos llamar base. No todos los pueblos primitivos usarían diez, o el número de dedos de ambas manos, como base. Algunos usarían sólo las dos manos en la contabilidad (y no los dedos de las manos), y su número máximo sería dos. Para otros, los dedos de una mano bastarían; su máximo sería cinco. Otros, inclusive, combinarían los dedos de ambas manos y los dedos de ambos pies; entonces su número de base sería el veinte.





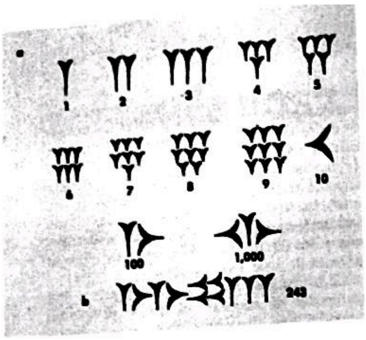
1. En los cálculos de los tiempos primitivos, el número uno podía representarse por un dedo o un palito (a). El número diez podía representarse por los diez dedos de ambas manos o por una piedra pequeña (b). Diez piedras pequeñas serían equivalentes a una piedra grande, que representaría cien (c). En este sistema, el número diez marcaba el número tope en la cuenta; en otras palabras, el sistema estaba basado en la decena.



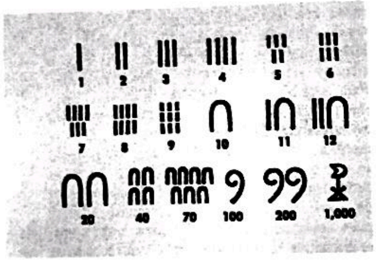
2. Si usáramos el método de cálculo ilustrado en la Figura 1, representaríamos el número 378 por tres piedras grandes (representando 300), siete piedras pequeñas (representando setenta) y ocho palitos (representando ocho).

Como las piedras, piedrecillas y palitos son incómodos de manejar, en cuanto el hombre aprendió a escribir creó símbolos que representaban los números. El símbolo empleado para escribir un número se denomina *guarismo*. Alrededor de 5.000 años atrás, los babilonios desarrollaron la escritura cuneiforme, o en forma de cuña, en la cual los diferentes símbolos se trazaban con una varilla de punta afilada sobre arcilla húmeda, moldeada en ladrillos lisos. Había símbolos especiales para uno, diez y cien; Y representaba 1; < , diez; Y> , cien. Para escribir una cifra, los babilonios repetían los símbolos disponiéndolos como fuera necesario. En la Figura 3*a* mostramos cómo se escribían los primeros diez números y cómo se escribían 100 y 1.000. La Figura 3*b* da el equivalente babilónico de 243. Dicho sistema de escritura numérica se denomina aditivo, pues debemos sumar los símbolos entre sí para encontrar el resultado total. Para representar los números más altos, los babilonios usaron un procedimiento multiplicativo. Por ejemplo, el símbolo <Y> representaba diez veces cien, o sea, mil. Si el símbolo de un número menor precedía al símbolo de un número mayor, indicaba que los números debían multiplicarse entre sí. Si el símbolo mayor estaba primero, los números debían sumarse.

Los antiguos egipcios poseían un sistema puramente aditivo –el sistema jeroglífico– que empleaba los símbolos dados en la Figura 4. La escritura numérica era un asunto engorroso. El número 527, por ejemplo, debería escribirse como está indicado en la Figura 5. Los egipcios usaban este sistema con propósitos decorativos, para monumentos de piedra, obe-



3. Números cuneiformes babilónicos; a: los símbolos para los números 1 a 10 y de 100 a 1.000; b: cómo podría escribirse el número 243.



4. El sistema numerico jeroglífico usado por los egipcios, particularmente con propósitos decorativos. Tenian otro sistema –el sistema hierático– para los cálculos de todos los días.

5. El sistema jeroglífico egipcio fue aditivo, y la escritura de números se volvió engorrosa. Arriba: cómo se escribiría el número 527.

liscos (Figura 6) y demás. También tenían un sistema numérico llamado hierático, que fue mucho más eficiente que el sistema jeroglífico y les servía en las contabilidades diarias.

Los antiguos griegos desarrollaron varios métodos de escritura de números. En la versión última y más ampliamente adoptada, usaban todas las letras de su alfabeto, más tres símbolos adicionales. Cada letra significaba un valor determinado. Los nueve primeros símbolos representaban los números de uno a nueve, los nueve siguientes, las decenas de diez a noventa; los últimos nueve, las centenas de cien a novecientos. No tenían símbolo alguno para el cero (Figura 7).

Si querían representar unidades de mil, los griegos agregaban una barra a la izquierda de las primeras nueve letras. Así **/Γ** representaba el 3.000; **/Z**, el 7.000. El número 4.627 se escribía **<u>/ΔXKZ</u>**: la línea horizontal de base indicaba qué letras formaban una cifra.

Los hebreos también usaron su alfabeto en la escritura de números (Figura 8).

Los antiguos chinos emplearon símbolos semejantes a barras para representar los números. Tenían un sistema de base, o sea, que un símbolo numérico tomaba diferentes valores de acuerdo con el lugar que ocupaba en el número escrito. En la Figura 9 damos los números de uno hasta nueve y de diez hasta noventa. Se advertirá que en los números de uno a nueve, cada barra vertical representaba una unidad y cada una de las horizontales era igual a cinco unidades. En los números de diez a noventa, cada barra horizontal representaba diez unidades, y cada barra vertical, cincuenta.

Las centenas se escribían del mismo modo que las unidades. Así, el símbolo II representaba tanto dos como doscientos, según su posición en el número. Las unidades de mil se escribían del mismo modo que las decenas, las decenas de mil, del mismo modo que las unidades, y así sucesivamente. El número 7.684 se escribía IIII ⊥T⊥ ; para el cero no había ningún símbolo: debía indicarse mediante un espacio. El número 7.004, por ejemplo, es ⊥ IIII . Si no se reconocía el espacio como tal, el número podría leerse como 74 en vez de 7.004.

Los romanos posiblemente derivaron su sistema numérico de los etruscos, los habitantes más primitivos de Italia: utilizaban letras para los números: "I" representaba uno, "V" cinco, "X" diez, "L" cincuenta, "C" cien, "D" quinientos, y "M" mil. Dos II representaban el número dos; tres I, el número tres; se colocaba una I antes de V para formar cuatro. Ésta fue la aplicación del principio llamado sustractivo. Si el símbolo de un número menor precedía al símbolo de un número mayor, el número menor debía restarse del mayor. VI representaba seis; VII, siete; VIII, ocho. El símbolo para el nueve era IX (principio sustractivo). Se muestran otros números en la Figura 10.

6. Los egipcios usaban los decorativos números jeroglíficos en monumentos tales como el típico obelisco ilustrado a la izquierda.

Así se procedía en el sistema numérico romano. Si un número era menor que el que lo seguía, aquél se restaba del segundo número; si era tan grande como el número siguiente, o mayor que él, el segundo número se sumaba al primero. Así LX era sesenta; XL, cuarenta.

Los números romanos se usan aún en varios idiomas occidentales para ciertos propósitos específicos. Algunas veces indican números de capítulo o de volumen. También se emplean en las esferas de algunos relojes y en monumentos conmemorativos y lápidas. Para nosotros, este sistema numérico parece extremadamente complicado. Por ejemplo, indicamos el número mil ochocientos ochenta y ocho con sólo cuatro símbolos: 1888. En números romanos se requerían trece símbolos, a saber: MDCCCLXXXVIII.

Nuestro propio sistema de números, compuesto por los llamados números arábigos, debería denominarse realmente indo-arábigo, dado que el sistema se originó en la India (no mucho antes de la era cristiana) y fue adoptado por los árabes después. Los árabes conquistaron una gran parte de España en el siglo VIII y en su oportunidad introdujeron los números indo-arábigos en la tierra conquistada. En la Figura 11 mostramos cómo aparecían éstos en el siglo X. El sistema fue gradualmente adoptado por otros pueblos de Europa. Para el siglo XV, los símbolos del sistema habían adquirido la forma que nos es tan familiar:

1, 2, 3, 4, 5, 6, 7, 8, 9, 0

La base de nuestro sistema es diez, y así se denomina sistema decimal (*decem* significa "diez" en latín). Es un sistema de situación,

en el que la posición de un símbolo indica su valor particular. Corriéndose un lugar a la izquierda, aumenta el valor diez veces, como se muestra en el siguiente diagrama:

10.000	1.000	100	10	1
(10 × 1.000)	(10 × 100)	(10 × 10)	(10 × 1)	

En el número 4.962, el 4 representa cuatro millares; el 9, nueve centenas; el 6, seis decenas; el 2, dos unidades. Por tanto, 4.962 es realmente 4.000 + 900 + 60 + 2. Tómese el número 7.004. El 7 representa siete millares; el primer cero, ninguna centena; el segundo cero, ninguna decena; el 4, cuatro unidades. Como hay un símbolo especial para el cero, no hay posibilidades de error en la lectura de dicho número, como lo habría si el cero se indicara, como en el primitivo sistema chino, dejando un espacio.

Con nuestros diez símbolos podemos escribir cualquier número, no importa cuán grande sea. Supongamos que empezamos con el número 1. Si colocamos un cero a la derecha de éste, el 1 está en la columna de las decenas e indica diez. Continuando la colocación de ceros a la derecha en esta forma, hacemos el valor de 1 diez veces mayor por cada cero que escribimos. En nuestro sistema también podemos escribir números extremadamente pequeños usando una coma decimal. El primer número después de una coma decimal indica el número de décimos; el segundo, el número de centésimos; el tercero, el número de milésimos y así sucesivamente.

0,000001 es un millonésimo; 0,00000001, un cien millonésimo. Podríamos hacer el número más y más pequeño aumentando el número de ceros entre la coma decimal y el 1.

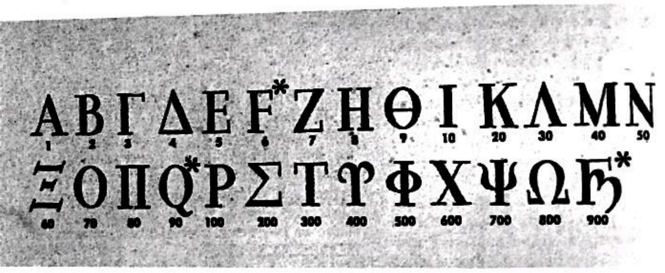
Por supuesto, números extremadamente grandes o extremadamente pequeños señalados en esta forma serían incómodos para escribir y usar. Mostramos en otra parte cómo indicar dichos números convenientemente.

No todos los sistemas numéricos necesitan tener diez como base. Varios otros se han usado o propuesto. Para mostrar las diferencias entre sistemas que tienen distintas bases, examinemos el siguiente párrafo:

"Un niño entró a primer grado a los diez años. Después de asistir a la escuela durante 22, ingresó a los estudios superiores a la edad de 32. Tenía 41 años de edad cuando se graduó, y se casó a la edad de 100. Murió a la edad de 240 años."

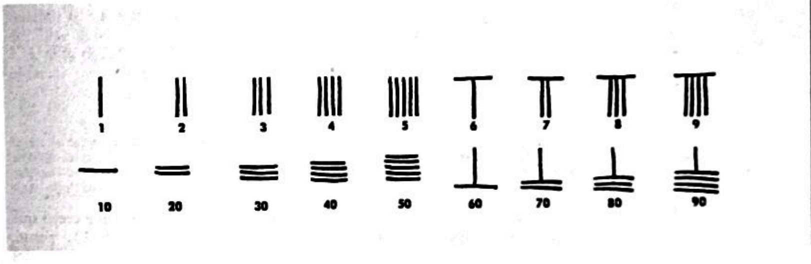
Si aplicamos nuestro propio sistema numérico con base diez, a este párrafo, no tendría mucho sentido. Sin duda, el niño en cuestión debe de haber vivido una vida muy lenta en realidad. Supongamos que usamos un sistema de base cinco para interpretar las cifras en el párrafo. En tal sistema, usaríamos sólo los símbolos 1, 2, 3, 4 y 0. Corriendo un espacio a la izquierda incrementaríamos el valor cinco veces. El último número en una cifra dada representaría el número de unidades; el penúltimo número, el número de cincos (5 × 1); el antepenúltimo, el número de veinticincos (5 × 5); el tercer número desde el último, el número de ciento veinticincos (5 × 25), etc.

Apliquemos este sistema a las diferentes cifras del pasaje de arriba. 10 = un cinco y ninguna unidad = cinco; 22 = dos cincos y dos unidades = doce; 32 = tres cincos y dos unidades = diecisiete; 41 = cuatro cincos y una unidad = veintiuno; 100 = un veinticinco, ningún cinco y ninguna unidad = veintincinco; 240 = dos veinticincos, cuatro cincos y ninguna unidad = setenta. Interpretado de esta forma, el párrafo es un enunciado perfectamente razonable.

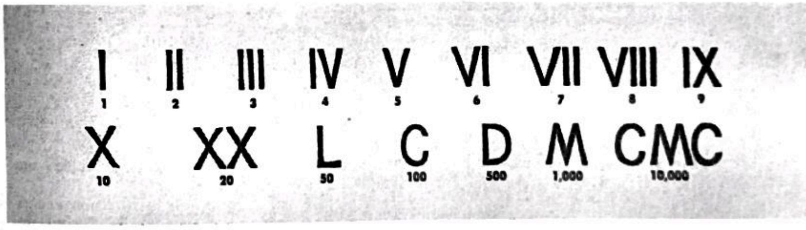


7. Los antiguos griegos escribían los números usando las letras de su alfabeto, más tres símbolos adicionales, indicados arriba con asteriscos.

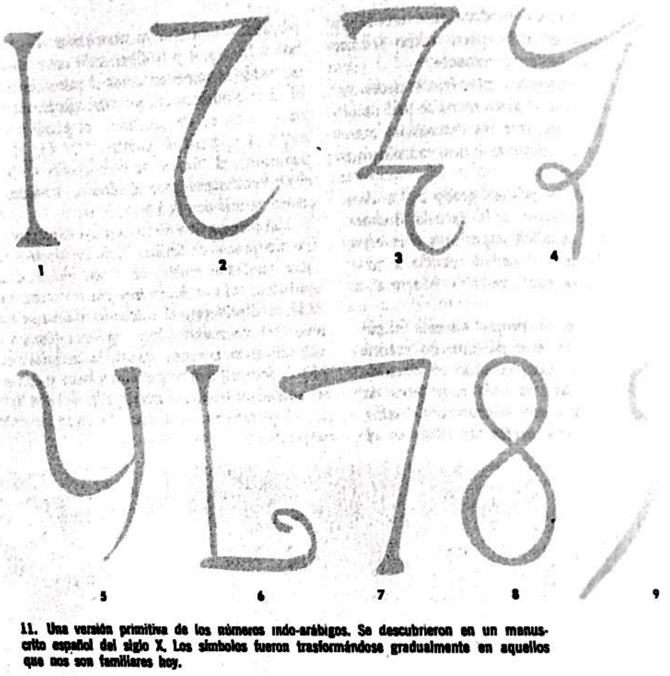
8. Como los griegos, los antiguos hebreos usaron las letras del alfabeto para escribir sus números.



9. Los simples símbolos como palotes (arriba) eran utilizados por los antiguos chinos en su sistema de escritura de números.



10. Los símbolos del sistema de numeración romano eran letras. Este sistema se usa aun para determinados propósitos, tales como numerar capítulos en algunos libros.



11. Una versión primitiva de los números indo-arábigos. Se descubrieron en un manuscrito español del siglo X. Los símbolos fueron trasformándose gradualmente en aquellos que nos son familiares hoy.

El sistema numérico denominado sistema binario tiene base dos. Utiliza sólo dos símbolos, 1 y 0. En el sistema binario, corriéndose un lugar hacia la izquierda se duplica el valor, como se indica en el siguiente diagrama:

32	16	8	4	2	1
(2 × 16)	(2 × 8)	(2 × 4)	(2 × 2)	(2 × 1)	

Veamos cómo se confirmaría el número 101101 en este esquema:

32	16	8	4	2	1
1	0	1	1	0	1

Como muestra el diagrama, 101101 en el sistema binario representa $(1 \times 32) + (0 \times 16) + (1 \times 8) + (1 \times 4) + (0 \times 2) + (1 \times 1) = 32 + 8 + 4 + 1 = 45$.

El sistema binario fue ardientemente defendido por el gran filósofo y matemático alemán del siglo XVII Gottfried Wilhem von Leibniz por su simplicidad y también porque creyó que reflejaba la creación. La unidad (1), pensó, representaba a Dios; el cero (0), el vacío del que se creaban todas las cosas.

El sistema binario tiene uso práctico real en física. También sirve en el dispositivo de cálculo denominado computadora electrónica o "cerebro electrónico". Este instrumento funciona con electricidad; la corriente se conecta o desconecta. Cuando está desconectada indica el "0" del sistema binario; cuando la corriente está conectada, indica el número "1" de este sistema. La computadora moderna puede realizar los cálculos más intrincados con sólo el "0" y el "1". En la foto de abajo se muestra una compleja computadora moderna.

Los indios mayas, que desarrollaron una civilización destacada en América Central durante los primeros siglos después de Cristo, usaban un sistema de base veinte. Hay indicios de este método de contabilidad en varios idiomas modernos, incluyendo el francés. Los números franceses, hasta el sesenta y nueve, corresponden a los nuestros. Pero setenta en francés es sesenta-diez (*soixante-dix*); setenta y uno es sesenta-once (*soixante et onze*), y así sucesivamente hasta que llegamos a ochenta. Ochenta es cuatro veintes (*quatre-vingts*). Después vienen cuatro veintes y uno (*quatre-vingt-un*); cuatro veintes y dos (*quatre-vingt-deux*). Noventa es cuatro veintes y diez (*quatre-vingt-dix*); noventa y uno es cuatro veintes y once (*quatre-vingt-onze*), y así sucesivamente, hasta cien. En inglés, también hay indicios de contar por veintes. En la versión de la Biblia del rey Jacobo, leemos que el lapso de vida humana es "tres veintes y diez". Abraham Lincoln empezó su inmortal discurso de Gettysburg del 19 de noviembre de 1863 con las palabras "cuatro veintenas y siete años atrás".

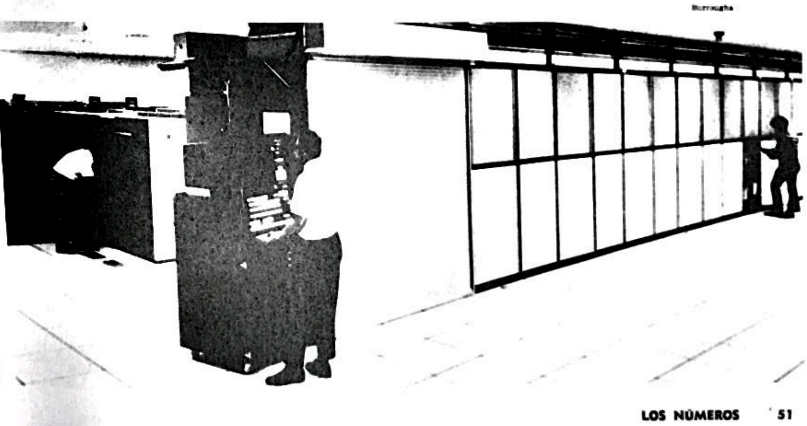
La Sociedad Duodecimal de América ha defendido la adopción del sistema duodecimal o de "base doce". En éste hay doce símbolos: 1, 2, 3, 4, 5, 6, 7, 8, 9, *t* (representa diez), *e* (representa once) y 0. Corriendo un lugar a la izquierda en el sistema duodecimal, se incrementa el valor doce veces, como sigue:

20.736	1.728	144	12	1
(12 × 1.728)	(12 × 144)	(12 × 12)	(12 × 1)	

El número 6*et*4 en el sistema duodecimal representa $(6 \times 1.728) + (11 \times 144) + (10 \times 12) + (4 \times 1)$. Si realizamos estas multiplicaciones y sumas, vemos que 6*et*4 es igual a 12.076.

Hay muy poca probabilidad de que, para los propósitos comunes de realizar operaciones, nuestro sistema decimal sea alguna vez reemplazado por otro con diferente base.

Abajo: la computadora ILLIAC IV y otros sistemas de computación avanzados que pueden procesar cientos de millones de instrucciones por segundo. El desarrollo de las computadoras –una consecuencia del desarrollo del sistema de numeración binario– ha cambiado significativamente muchos aspectos de la vida moderna.



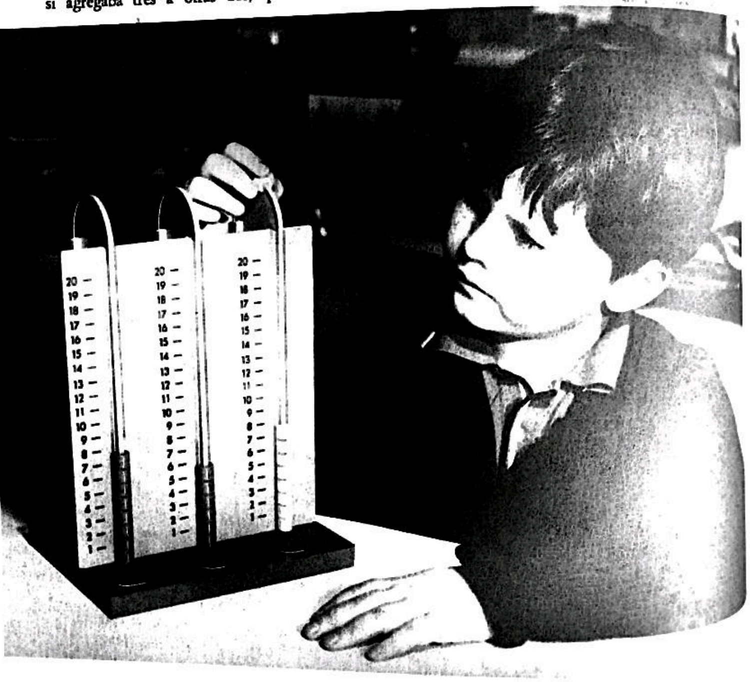
ARITMÉTICA

Los cálculos llevados a cabo con los números del sistema decimal constituyen el ramo de las matemáticas denominado aritmética. Hay seis operaciones fundamentales: adición, sustracción, multiplicación, división, potenciación (obtención de las potencias de los números) y radicación (extracción de las raíces de éstos).

ADICIÓN

La adición o suma representa el agrupamiento o reunión de números. Cuando un hombre primitivo quería resolver cuántas pieles tendría si agregaba tres a otras dos, quizá tendería tres pieles en el suelo, colocaría dos más, y entonces contaría el número total. Después de resolver este problema particular una y otra vez, ya no tendría que tender las pieles en el suelo; recordaría que siempre que agregara tres a otras dos, el resultado sería cinco pieles, y realizaría la suma $3 + 2 = 5$ mentalmente. Del mismo modo, llegaría a entender que $1 + 1 = 2$, $1 + 2 = 3$, $1 + 3 = 4$, etc.

Cuando se es capaz de sumar hasta $9 + 9$ se puede realizar cualquier otra suma. Supongamos que el problema es $23 + 49$. Recordemos que 23 significa 2 decenas + 3 unidades y que [illegible]



significa 4 decenas + 9 unidades. Podemos plantear nuestro problema como sigue:

2 decenas + 3 unidades
+ 4 decenas + 9 unidades

Primero sumaríamos las unidades juntas: $9 + 3 = 12$. 12 es realmente igual a 1 decena más 2 unidades. Manteniendo el 2 en la columna de las unidades, sumaríamos el 1 a los otros números en las columnas de las decenas, dando $2 + 4 + 1$, ó 7 decenas. La respuesta entonces es 7 decenas + 2 unidades, o sea, 72.

SUSTRACCIÓN

En la sustracción o resta tomamos uno o más objetos de otro grupo de objetos. Supongamos que nuestro matemático primitivo tenía nueve pieles y quería saber cuántas pieles le quedaban si quitaba cinco; tendería las nueve pieles en el suelo, quitaría cinco pieles, y contaría las pieles que quedaban. Luego, llegaría a darse cuenta que quitar cinco de nueve siempre deja cuatro, y realizaría la resta mentalmente; extendería entonces este tipo de cálculo a otros números. Llegaría a saber, por ejemplo, que $9 - 4 = 5$ y que $7 - 6 = 1$. Necesitaría saber sólo unos pocos de dichos cálculos para poder restar cualquier cantidad de cualquier otra.

MULTIPLICACIÓN

La multiplicación es realmente una forma de suma. Si no supiéramos cómo multiplicar, podríamos encontrar la respuesta al problema: 5×7 por simple suma, porque 5×7 significa lo mismo que cinco sietes, ó $7 + 7 + 7 + 7 + 7$. Sumando los cinco sietes juntos, tendríamos 35. Más tarde o más temprano llegaríamos probablemente a darnos cuenta de que cuando resolvemos el problema de sumar cinco sietes: $7 + 7 + 7 + 7 + 7$, el resultado es siempre 35. Memorizaríamos esta operación particular y otras similares tales como tres ochos (3×8) y sietes nueves (7×9). Estas operaciones, hasta incluir 12×12, se encuentran en la muy conocida tabla de multiplicar. Si lográramos hallar el resultado de 9×9, podríamos realizar cualquier problema de multiplicación.

Por supuesto, todo esto requiere cierta cantidad de memorización. Algunos acostumbraban eludirla utilizando un método de multiplicación, denominado ***duplicación***, que incluye sólo multiplicación por dos y suma. Así opera la duplicación. Supongamos que quisiéramos resolver el producto de 24 y 18 (24×18); realizaríamos nuestros cálculos así:

(a)	*(b)*	*(c)*	*(d)*
1	24		
2	48	2	48
4	96		
8	192		
16	384	16	384
			432

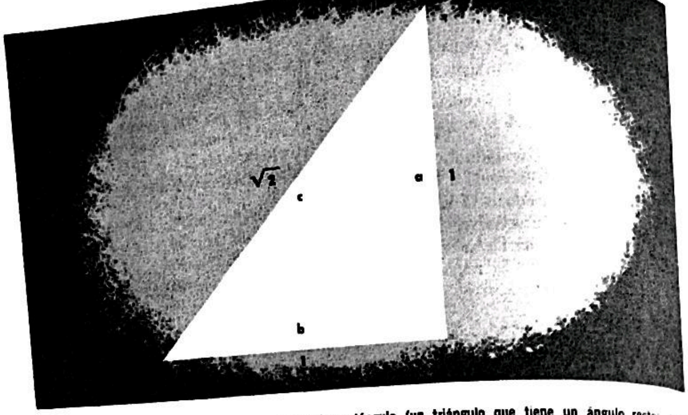
432, en la última columna, es la solución de nuestro problema. Todo esto parece bastante complicado, pero es realmente sencillo.

Como se ve, fijamos cuatro columnas (*a*), (*b*), (*c*) y (*d*). El primer número en la columna (*a*) es siempre 1; el primer número en (*b*) es siempre el ***multiplicando***, que en este caso es 24. El multiplicando es un número que debe multiplicarse por otro número, conocido como ***multiplicador***. Continuamos multiplicando los números en la columna (*a*) por 2 hasta que llegamos a un número que sea justamente igual o menor que el multiplicador, que es 18. Nos detenemos en 16, pues 2 veces 16, ó 32, es mayor que 18. Ahora duplicamos el multiplicando en la columna (*b*), dando 48. Duplicamos 48, y continuamos duplicando los números en esta columna hasta que haya tantos números en la columna (*b*) como en la (*a*). Luego, seleccionamos de la columna (*a*) los números que, sumados, sean iguales al multiplicador, 18. Los números que seleccionamos son 2 y 16. Colocamos estos números en la columna (*c*) en los mismos lugares que ocupan en la columna (*a*). En la columna (*d*) ponemos los números de la columna (*b*) que están al lado de 2 y 16 de la columna (*a*). La suma de los números en la columna (*d*), 432, es la solución del problema 24×18. Por supuesto, cualquiera que sepa la tabla de multiplicar podría multiplicar 24 por 18 en mucho menos tiempo del que le llevaría resolver el problema por medio de la duplicación.

DIVISIÓN

La división es un tipo de resta. Si dividimos 12 por 4, queremos saber cuántas veces 4 está

Un niño utilizando un ábaco de anillos. El ábaco es un dispositivo que se ha empleado durante miles de años para resolver cálculos aritméticos sencillos.



Esta figura es un triángulo rectángulo (un triángulo que tiene un ángulo recto). Lado $a = 1$; lado $b = 1$. El cuadrado del lado c es igual a la suma de los cuadrados de los otros dos lados. Por tanto, $c^2 = 1 + 1 = 2$. Si $c^2 = 2$, $c = \sqrt{2}$.

contenido en 12. Efectuamos la siguiente serie de restas:

$$\begin{array}{r}12\\-4\\\hline 8\end{array}\qquad\begin{array}{r}8\\-4\\\hline 4\end{array}\qquad\begin{array}{r}4\\-4\\\hline 0\end{array}$$

Hemos restado 4 de 12; luego 4 del 8 restante; luego 4 del 4 restante. Queda cero. Usamos el 4 como sustraendo tres veces. En consecuencia, la respuesta al problema 12 dividido por 4, ó $12 \div 4$, es 3.

POTENCIACIÓN

En el proceso denominado potenciación elevamos un número a cualquier potencia deseada. El número que debe elevarse a la potencia en cuestión se denomina base. Para elevar 2 a la tercera potencia, multiplicamos repetidamente la base por sí misma, tres veces, así: $2 \times 2 \times 2$. Podemos indicar la potencia colocando una pequeña cifra, denominada exponente, a la derecha de la base y arriba de ésta. Por ejemplo, 2 a la tercera potencia, ó $2 \times 2 \times 2$, se escribe 2^3.

Si el exponente es 1, indica que el número no está elevado a una potencia más alta, sino que permanece inalterado. Por tanto, $2^1 = 2$; $5^1 = 5$. Cuando usamos el exponente cero, mostramos que la base debe dividirse por sí misma. Cualquier número con el exponente 0 es siempre igual a 1. Por ejemplo, $4^0 = 4 \div 4 = 1$; $10^0 = 10 \div 10 = 1$. Hay también exponentes negativos. Una base con un exponente negativo es igual a la recíproca de la base $\left(\frac{1}{\text{base}}\right)$ con el correspondiente exponente positivo. Así $5^{-1} = \frac{1}{5^1} = \frac{1}{5}$; $3^{-2} = \frac{1}{3^2} = \frac{1}{9}$

RADICACIÓN

En el proceso denominado radicación hallamos las raíces de los números. Dado determinado número tratamos de encontrar otro que multiplicado por sí mismo el número de veces deseado nos dé el primero. El número que se multiplica por sí mismo se denomina raíz. Si queremos encontrar qué número multiplicado por sí mismo dará 4, decimos que estamos tratando de hallar la raíz cuadrada de 4, que se escribe $\sqrt{4}$. El número 2, multiplicado por sí mismo, o sea, usado como factor dos veces, da 4. Por tanto, 2 es la raíz cuadrada de 4. La raíz cuadrada de cuatro puede ser también -2, pues $-2 \times -2 = 4$. La raíz cúbica del número 8, que se escribe $\sqrt[3]{8}$, es el número que usado como factor tres veces, da 8; $2 \times 2 \times 2 = 8$; por tanto, $\sqrt[3]{8} = 2$.

FRACCIONES

Los números tales como 1, 2, 3, 5, 10, 120 y 3.000 se denominan *números enteros* o *enteros*. Cuando uno multiplica un entero por otro, el resultado es siempre un entero: $5 \times 6 = 30$.

$7 \times 9 = 63$. Sin embargo, no siempre es posible obtener un entero como resultado cuando se divide un entero por otro. Cierto es que si dividimos 8 manzanas en 4 partes iguales, cada parte contendrá 2 manzanas. Pero si 8 manzanas deben dividirse en 3 partes iguales, el resultado no será un entero. Pues, $8 \div 3 = \frac{8}{3}$, que es una fracción, o "número quebrado".

Esto no significa que tengamos que dividir todas nuestras 8 manzanas en tercios y dar a cada persona 8 tercios. Observamos que $\frac{3}{3}$ es exactamente equivalente a una manzana entera y que $\frac{6}{3}$ es exactamente equivalente a dos manzanas enteras. Para dividir 8 manzanas en 3 partes iguales, damos a cada persona 2 manzanas. Dividimos las 2 manzanas restantes en tercios y damos a cada persona 2 de los tercios. Otra forma de plantear esto es que $8 \div 3 = 2\frac{2}{3}$; $2\frac{2}{3}$ representa la suma de un entero y una fracción, y se denomina *número mixto*.

En una fracción tal como $\frac{2}{5}$, el número colocado encima de la línea se denomina *numerador*, el número debajo de la línea es el *denominador*. El matemático no considera $\frac{2}{5}$ como dos números, sino como un número único de valor definido. Si escribimos $\frac{2}{5}$ en la forma decimal, 0,4, vemos que la fracción es realmente un número único.

Cada fracción es un cociente compuesto de dos enteros. Es posible trascribir todas las fracciones usando el siguiente esquema. Primero colocamos las fracciones con denominador 1, luego aquellas de denominador 2, luego las de denominador 3, como sigue:

Denominador 1 .... $\frac{1}{1}, \frac{2}{1}, \frac{3}{1}, \frac{4}{1}, \frac{5}{1}$, etc.

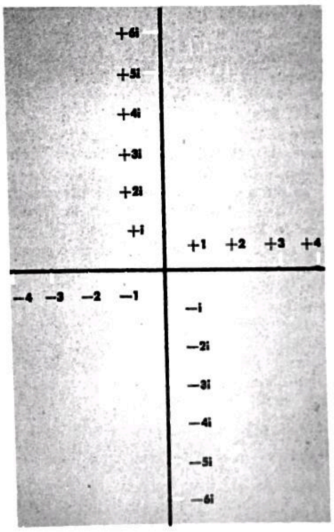
Denominador 2 .... $\frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \frac{4}{2}, \frac{5}{2}$, etc.

Denominador 3 .... $\frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \frac{4}{3}, \frac{5}{3}$, etc.

Evidentemente, llevaría una eternidad trascribir todas las fracciones de esta forma. Sin embargo, suponiendo que dicha lista de fracciones pudiera llegar a completarse, cualquier fracción podría encontrar un lugar en ella.

Las fracciones pueden sumarse, restarse, multiplicarse y dividirse según las reglas comunes de la aritmética. El resultado será siempre un entero, o una fracción (compuesta de dos enteros), o un número mixto (compuesto de un entero más una fracción). Al dividir un número por una fracción usamos la regla de "invertir el divisor y multiplicar". Así $\frac{2}{5} \div \frac{7}{0} = \frac{2}{5} \times \frac{9}{7} = \frac{18}{35}$.

En este diagrama, los números arábigos positivos representan unidades hacia el este; los números arábigos negativos, unidades hacia el oeste; las íes positivas, unidades hacia el norte; las íes negativas, unidades hacia el sur. *i* es el símbolo para $\sqrt{-1}$.



En una fracción tal como $\frac{18}{35}$, el denominador, 35, es realmente un divisor, pues $\frac{18}{35} = 18 \div 35$.

Ahora bien: en matemáticas, el cero es siempre excluido como divisor, y por tanto nunca puede ser un denominador. Se hace esto para evitar resultados absurdos. Por ejemplo, si debiéramos admitir la división por cero, podríamos probar que $6 = 1$. Sería así: sabemos que $6 \times 0 = 0$. Si dividimos cada lado de la ecuación por cero, tenemos $\frac{6 \times 0}{0} = \frac{0}{0}$, que podría escribirse como $6 \times \frac{0}{0} = \frac{0}{0}$. Ahora bien, cualquier número dividido por sí mismo es igual a 1. Por tanto, podríamos cambiar la ecuación anterior para leer $6 \times 1 = 1$ ó $6 = 1$. En realidad si admitiéramos la división por cero, podríamos probar prácticamente cualquier cosa.

ENTEROS NEGATIVOS

Si volvemos ahora a los enteros, advertimos que cuando restamos un entero positivo de otro, el resultado no siempre es positivo. Es cierto que $7 - 4 = +3$, que generalmente se escribe simplemente como 3. Pero supongamos que queremos restar 7 de 4. Para hacer posible la resta $4 - 7$ inventamos un nuevo tipo de número, denominado entero negativo, que se representa por un entero con el signo menos adelante. El resultado para la resta de arriba, $4 - 7$, sería -3.

La mecánica de resolver problemas tales como $12 - 6$, $9 - 7$, $6 - 13$, etc., es bastante sencilla. Restamos el número menor del mayor y le damos al resultado el signo (más o menos) del número mayor. A modo de ejemplo, $12 - 6 = 6$; $9 - 7 = 2$; $4 - 1 = 3$; $4 - 9 = -5$; $3 - 7 = -4$; $6 - 13 = -7$.

Hay tantos enteros negativos posibles como enteros positivos. Comenzando con 0, podemos construir una lista de enteros negativos a la izquierda de 0 y una lista de enteros positivos a la derecha de 0, como sigue:

$$\ldots -4, -3, -2, -1, 0, +1, +2, +3, +4 \ldots$$

Podríamos extender en esta forma indefinidamente la lista de los enteros negativos y positivos. Hay fracciones negativas así como enteros negativos. Cada fracción positiva, tal como $\frac{3}{5}$ tiene una equivalente negativa, tal como $-\frac{3}{5}$.

El conjunto completo de números que hemos estado tratando hasta ahora –enteros positivos, enteros negativos, fracciones positivas, fracciones negativas, y cero– se denomina *sistema de números racionales*. Podemos definir un número racional como cero, o un entero (positivo o negativo), o una fracción (positiva o negativa), cuyo numerador y denominador sean ambos enteros. 5 es un número racional; también lo son -8, $\frac{1}{2}$, $-\frac{3}{4}$.

Se han ideado reglas para la suma, resta, multiplicación y división de números racionales positivos y negativos; por tanto, no se presentarán resultados ilógicos. Una de estas reglas es: "El producto de dos números del mismo signo es positivo". Por ejemplo, $4 \times 4 = 16$; asimismo $-4 \times -4 = 16$. Otra de estas reglas es: "El producto de dos números de distinto signo es negativo". Ejemplos de la regla son $4 \times -4 = -16$; $-7 \times 6 = -42$.

NÚMEROS IRRACIONALES

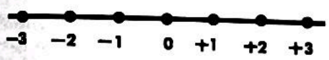
La raíz cuadrada de un número, como hemos visto, es uno de sus dos factores iguales. Así, la raíz cuadrada de 25, ó $\sqrt{25}$, es 5, pues $5 \times 5 = 25$. Pero la raíz cuadrada de 2, ó $\sqrt{2}$, no puede expresarse por ningún número racional. Sabemos que $\sqrt{2}$ debe estar entre 1,4 y 1,5 pues $1{,}4^2 = 1{,}96$ y $1{,}5^2 = 2{,}25$. Podríamos aproximarnos a la raíz cuadrada de 2 tomando más y más decimales. Así $1{,}41^2 = 1{,}9881$ y $1{,}414^2 = 1{,}999396$. Pero aun cuando usáramos una raíz con una docena de decimales nunca podríamos llegar a un número racional cuyo cuadrado fuera 2.

Sin embargo, en la labor matemática debemos tratar con números tales como $\sqrt{2}$. Por ejemplo, uno de los teoremas más conocidos de geometría es que en cualquier triángulo rectángulo (un triángulo que tiene un ángulo recto), el cuadrado de la hipotenusa (el lado opuesto al ángulo recto) es igual a la suma

de los cuadrados de los otros dos lados. La Figura 1 representa un triángulo rectángulo, en el que los lados a y b son ambos iguales a 1; c es la hipotenusa. c^2, por tanto, es igual a $a^2 + b^2$ y como $a^2 = 1$ y $b^2 = 1$, $c^2 = 2$. c, por tanto, debe ser igual a $\sqrt{2}$. Evidentemente, $\sqrt{2}$ es un número real, que obtenemos por un correcto cálculo matemático, pero, como hemos visto, no es un número racional. Lo llamamos un *número irracional*.

Hay una gran cantidad de números irracionales además de $\sqrt{2}$. La raíz cuadrada de 3 ($\sqrt{3}$) es un número irracional. Como también la raíz cuadrada de 5 ($\sqrt{5}$) y la raíz cúbica de 7 ($\sqrt[3]{7}$). Todos éstos son números reales, aun cuando no podamos expresarlos por enteros ni por cocientes compuestos de enteros.

Si combinamos todos los números irracionales con todos los números racionales, obtenemos un conjunto muy grande de números denominado *sistema de números reales*. Todos los números del sistema pueden representarse por puntos de una línea recta. Hagamos que un punto represente el número 0. Luego, hagamos que los puntos de la derecha representen los enteros positivos y aquellos de la izquierda, los enteros negativos, como sigue:



Si el espacio entre los números mostrados arriba se subdivide en tantas partes como sea posible, cada uno de los puntos de subdivisión representará un número racional. No importa cuántas divisiones podamos hacer en esta forma; siempre quedarán vacíos. Si rellenamos los vacíos con puntos que representen todos los números irracionales posibles, tales como $\sqrt{2}$, $\sqrt{5}$, etc., la línea quedará completamente llena. Esta línea se conoce como el eje de los números reales, o *cantidad continua*.

NÚMEROS IMAGINARIOS

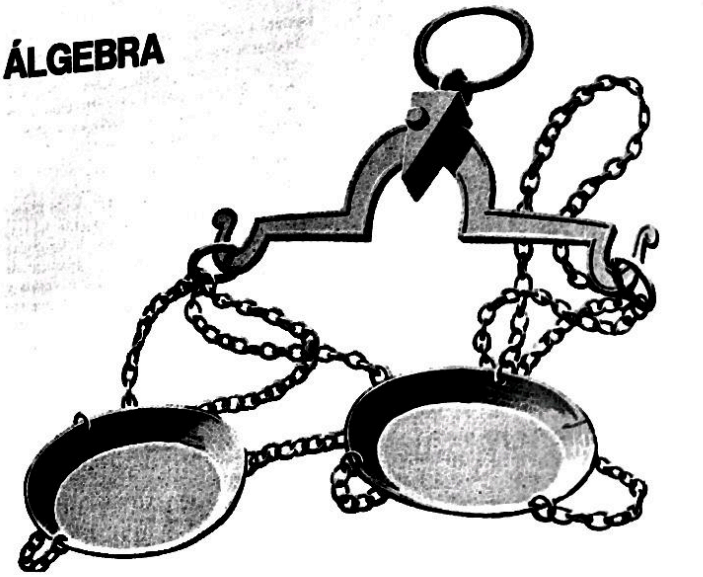
Hay aun otros números además de los números reales. Consideremos el número $\sqrt{-1}$. Esto parece ser una contradicción de las reglas, pues un cuadrado es siempre el producto de dos números iguales con los mismos signos y es siempre positivo. Por tanto, ningún número multiplicado por sí mismo puede dar un número real negativo, y parecería inútil tratar de obtener la raíz cuadrada de tal número. Sin embargo, los matemáticos han usado el número $\sqrt{-1}$ para formar la base de un sistema numérico denominado *números imaginarios*, o *números complejos*. $\sqrt{-1}$ se indica a menudo con el símbolo i.

Los números complejos se usan en física para representar fuerzas. Supongamos que trazamos dos líneas que se cortan en ángulos rectos, tal como en la Figura 2. Usamos 1 para representar una unidad hacia el este y -1 para representar una unidad en la dirección opuesta, o sea el oeste. Por tanto, podemos interpretar $+i$ $(+\sqrt{-1})$ como unidad hacia el norte de la línea este-oeste, y $-i$ $(-\sqrt{-1})$ como una unidad hacia el sur. La notación $5 + 3i$ significaría: "Vaya 5 unidades hacia el este y 3 unidades hacia el norte". $-3 + (-2i)$ significaría: "Vaya 3 unidades hacia el oeste y 2 unidades hacia el sur".

Se ha propuesto un nuevo sistema de numeración que sería útil en matemáticas superiores. En este sistema, conocido como numeración no normalizada, el intervalo entre cada número real ordinario puede verse como compuesto de una serie de números no normalizados. Los números no normalizados, considerados desde dentro del intervalo, pueden pensarse como números reales ordinarios; desde fuera del intervalo, la serie de números es infinitamente pequeña (izquierda) o infinitamente grande (derecha).

©1976 by The New York Times Company, reimpreso con autorización

ÁLGEBRA



En el artículo "Aritmética" estábamos interesados en números particulares, que se expresaban por símbolos. "Sesenta y siete" es un número particular. Para resolver problemas aritméticos en los que sesenta y siete toma parte, usamos los símbolos "6" y "7", combinados como 67. Ahora vamos a considerar un ramo de las matemáticas en el que un símbolo, tal como una letra –*a*, *b* o *c*–, no representa un número particular, sino una clase completa de números. Este tipo de matemáticas se denomina *álgebra*.

Podemos ilustrar la diferencia entre aritmética y álgebra por un ejemplo muy sencillo. Tomemos el número 4. Multipliquémoslo por 5 ($4 \times 5 = 20$); sumemos 4 ($20 + 4 = 24$); dupliquemos el resultado ($24 \times 2 = 48$); restemos 8 ($48 - 8 = 40$); dividamos por el número original, 4 ($40 \div 4 = 10$). El resultado de todas estas operaciones, como se ve, es 10. Para llegar al resultado final, 10, hemos recurrido al método de la aritmética, usando números particulares, de principio a fin.

Supongamos ahora que pensamos en cualquier número. Indiquemos "cualquier número" por el símbolo x y realicemos las mismas operaciones que antes. Multiplicamos x por 5 ($5 \times x$, se escribe $5x$); sumamos 4 ($5x + 4$), duplicamos el resultado ($10x + 8$); restamos 8 ($10x$); dividimos por el número original que es x $\left(\frac{10x}{x} = 10\right)$. El resultado, como se ve, es 10. Aquí hemos estado usando los métodos del álgebra, porque x puede reemplazarse por cualquier número. Podríamos sustituirlo por 2, 3 ó 15, y el resultado final sería 10.

Cuando un número generalizado, representado por una letra (tal como a), se multiplica por un número particular (tal como 5) o por otro número generalizado (tal como b), no usamos signos de multiplicación, pero indicamos la multiplicación colocando estos símbolos uno al lado del otro. Así $a \times b = ab$; $5 \times a = 5a$; $5 \times a \times b = 5ab$. No podríamos indicar la multiplicación de dos números particulares en esta forma: 7×5 no puede escribirse 75, porque 75 realmente representa $70 + 5$.

Consideremos otro ejemplo. En la ecuación $(2 + 3)^2 = 25$ estamos tratando con los números particulares 2 y 3, y el resultado es siempre 25. Pero supongamos que en vez de dos números particulares usamos la letra a para representar cualquier número, y la letra b, para representar cualquier otro número distinto de a. Tendríamos entonces $(a + b)^2 = a^2 + 2ab + b^2$.

Lo importante de $(a + b)^2 = a^2 + 2ab + b^2$ es que indica una relación general que se mantiene cierta para gran cantidad de números particulares. Si sustituyéramos 3 por a y 2 por b, tendríamos $(3 + 2)^2 = 3^2 + (2 \times 3 \times 2) + 2^2 = 9 + 12 + 4 = 25$. O podríamos sustituir 5 por a y 6 por b, dando $(5 + 6)^2 = 5^2 + (2 \times 5 \times 6) + 6^2 = 25 + 60 + 36 = 121$. El álgebra, la matemática de "números cualesquiera" o "variables", va al corazón mismo de la relación entre números. En general, puede decirse que tiene que ver con los números particulares sólo en cuanto a que son aplicaciones de principios generales. También se usa en la solución de ciertos problemas determinados en los que partimos de una o más cantidades desconocidas cuyos valores se indican por símbolos algebraicos.

UNA ANTIGUA DISCIPLINA

El estudio del álgebra se remonta a la antigüedad. Descubrimientos recientes han mostrado que los babilonios resolvieron problemas algebraicos, a pesar de que no tenían símbolos para las variables. Usaron sólo palabras para indicar tales números, y por esa razón su álgebra ha sido llamada álgebra retórica. El papiro Ahmes, un rótulo o rollo egipcio que data de 1600 a.C., tiene un número de problemas algebraicos en los que la incógnita se designa como un *hau*, que significa "un montón".

Poco progreso se realizó posteriormente en álgebra hasta que llegamos a Diofanto, un matemático griego que vivió en el siglo III d.C. Redujo los problemas a ecuaciones, representando la cantidad desconocida por un símbolo parecido a la letra griega Σ (sigma). También implantó un interesante sistema de abreviaturas en las que usaba sólo las letras iniciales de palabras y omitía todas las palabras innecesarias. Si debiéramos usar el método de Diofanto para presentar el problema: "Una incógnita elevada al cuadrado menos la incógnita dará veinte", expresaríamos primero el problema como "incógnita al cuadrado menos incógnita igual a veinte". Luego usaríamos las iniciales de todas las palabras, excepto la última, que la expresaríamos en cifras, de esta manera: I A C M I I 20.

En el siglo XVI, François Viète, un matemático francés, usó las vocales *a*, *e*, *i*, *o*, *u* para representar incógnitas, y las consonantes *b*, *c*, *d*, *f*, *g*, etc., para representar valores que se mantenían fijos a lo largo de un problema dado. René Descartes, gran filósofo francés del siglo XVII, propuso el sistema de símbolos algebraicos actualmente en uso. En este sistema, *a*, *b*, y *c* y otras letras próximas al comienzo del abecedario representan los números fijos. Las últimas letras del abecedario (*x*, *y*, *z*, y a veces también *w*) representan las incógnitas en un problema. Tan pronto como este simbolismo se hizo de uso general, el álgebra se trasformó muy rápidamente en un conjunto sistemático de reglas y teoremas que podrían aplicarse a todos los números.

La palabra "álgebra" se originó por el título de un trabajo sobre álgebra de un persa, Abú Abdallah Mohamed-ben-Mua Al-Khuwarizmi, que vivió en el siglo IX d.C. Escribió en árabe un trabajo denominado *Al-Jebr W'al Muqabala*, que significa "reintegración y reducción". Por *al-jebr*, o reintegración, se entendía la trasposición de términos negativos al otro miembro de una ecuación para hacerlos positivos. Cuando los árabes llegaron a España, trajeron esta palabra con ellos. Con el trascurso del tiempo, *al-jebr* se convirtió en "álgebra", y la palabra comenzó a aplicarse no sólo a una única operación, sino a las muchas operaciones comprendidas en álgebra.

TRES LEYES FUNDAMENTALES

El álgebra generaliza –o sea, expresa en términos generales– ciertas leyes básicas que rigen la suma, resta, multiplicación y división de todos los números.

(1) Cuando sumamos o multiplicamos dos enteros, el orden en que lo hagamos es indistinto. Así $2 + 3$ es lo mismo que $3 + 2$, y 4×3 es lo mismo que 3×4. Como esto es cierto para todos los enteros, establecemos las siguientes fórmulas algebraicas:

$$a + b = b + a; \quad ab = ba$$

Se denominan ***leyes conmutativas*** de la suma y la multiplicación.

(2) Cuando se suman o multiplican más de dos números, podemos agruparlos en cualquier orden que elijamos, y el resultado siempre será el mismo. Si se suma 2 a $(3+6)$, el resultado es el mismo que si sumamos $(2+3)$ a 6. Del mismo modo, dos veces el producto de 3×6 es lo mismo que tres veces el producto de 6×2.

Estos resultados se indican en las fórmulas:

$$a+(b+c)=(a+b)+c; \quad a(bc)=(ab)c$$

Éstas son las ***leyes asociativas*** de la suma y la multiplicación.

(3) Si un multiplicando tiene dos o más términos, un multiplicador debe operar sucesivamente sobre cada uno de ellos. Supongamos que deseamos multiplicar $3+2$ por 5, lo que podríamos expresar como $5(3+2)$. Multiplicaríamos primero 3 por 5 y luego 2 por cinco, dando $15+10$, o sea 25.

Esta regla se denomina ***ley distributiva*** de la multiplicación y se expresa por la fórmula siguiente:

$$a(b+c)=ab+ac$$

Supongamos que queremos multiplicar $a+b$ por $a+b$, que por supuesto sería lo mismo que $(a+b)^2$. Podríamos expresar el problema como sigue:

$$\begin{array}{r} a+b \\ \underline{a+b} \end{array}$$

De acuerdo con la ley distributiva de la multiplicación, multiplicamos primero a y b de la línea superior, por b de la línea de abajo. Luego multiplicamos a y b de la línea superior por a de la línea de abajo. Finalmente, sumamos los resultados:

$$\begin{array}{r} a+b \\ \underline{a+b} \\ ab+b^2 \\ \underline{a^2+ab} \\ a^2+2ab+b^2 \end{array}$$

Empleamos la misma ley distributiva de la multiplicación cuando multiplicamos 25 por 25. Si lo hacemos en la forma común, presentaríamos nuestros cálculos como sigue:

$$(\mathrm{I}) \qquad \begin{array}{r} 25 \\ \underline{25} \\ 125 \\ \underline{50} \\ 625 \end{array}$$

De hecho, como el 2 de 25 significa realmente 20, el problema es, pues:

$$\begin{array}{r} 20+5 \\ \underline{20+5} \end{array}$$

Usando la ley distributiva, multiplicamos: 5 en la primera línea por 5 en la segunda y 20 en la primera línea por 5 en la segunda. Luego, 5 en la primera línea por 20 en la segunda, y 20 en la primera línea por 20 en la segunda. Finalmente, sumamos los productos. Podríamos indicar estas operaciones como sigue:

$$(\mathrm{II}) \qquad \begin{array}{r} 20+\ 5 \\ \underline{20+\ 5} \\ 100+25 \\ \underline{400+100} \\ 400+200+25=625 \end{array}$$

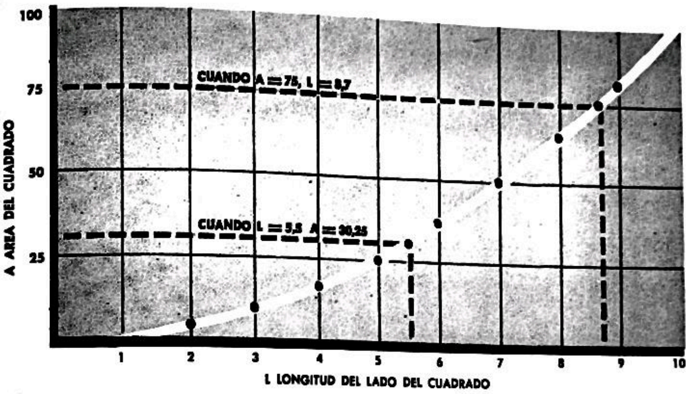
Si se comparan cuidadosamente (I) y (II) arriba se verá que básicamente se realizaron las mismas operaciones en ambos. Simplemente, recuérdese que, por ejemplo, cuando en (I) multiplicamos el 2 de la segunda línea por el 5 de la primera línea, aquél representa 20.

FÓRMULAS, TABLAS Y GRÁFICOS

Hay diferentes formas de mostrar cómo se relacionan diferentes cantidades. Podemos usar una fórmula, construir una tabla de valores o trazar un gráfico.

Tomemos la regla: "El área de un cuadrado es igual al cuadrado de la longitud de un lado". Ésta es una forma bastante vaga de expresar la relación en cuestión. Podríamos expresarla mucho más simplemente usando el símbolo A para representar el número de unidades cuadradas en el área, y el símbolo l para representar el número de unidades en el lado. La regla anterior puede, por tanto, expresarse como $A = l^2$. Esta regla abreviada se denomina fórmula, de la palabra latina que significa "forma pequeña".

Si la longitud del lado es 6, podemos obtener A, el área, sustituyendo 6 por l en la fórmula. A, entonces, resulta 6^2, o sea 36. Como se ve, hemos dado uso práctico a nuestra fórmula. Técnicamente hablando, hemos hallado el



1. Ilustración de la gráfica del área de un cuadrado, usando dos ejes: uno horizontal y otro vertical, de acuerdo con el método descrito en el texto. La gráfica muestra cómo aumenta el área de un cuadrado a medida que aumenta la longitud del lado. Es una forma de expresar la relación entre el área y la longitud del lado.

valor numérico de la fórmula para el valor 6.

También podemos representar la relación $A = l^2$ construyendo una tabla de valores. Supongamos que el lado del cuadrado es igual a 1; A es entonces 1^2 ó 1. Si el lado del cuadrado es igual a 2, $A = 2^2$, o sea 4. Sustituyendo sucesivamente l por los valores de 1 a 10, obtenemos la siguiente tabla de valores:

l	1	2	3	4	5	6	7	8	9	10
A	1	4	9	16	25	36	49	64	81	100

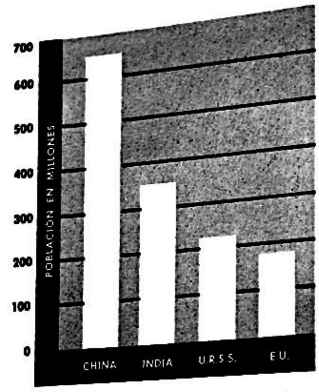
Esta tabla nos dice que si $l = 1$, $A = 1$; si $l = 2$, $A = 4$; si $l = 3$, $A = 9$, etc.

Hay aún otra forma de representar el área de un cuadrado. Podríamos construir un gráfico, como en la Figura 1. Primero trazamos dos líneas, denominadas ejes, que son perpendiculares entre sí. A lo largo del eje horizontal, o base, marcamos una serie de números a intervalos iguales, correspondientes a los valores de l en la tabla. Marcamos otra serie de números en el eje vertical. Estos números corresponden a los valores de A. En cada uno de los valores de l, 1, 2, 3, 4, etc., elevamos una perpendicular. Después marcamos el valor apropiado de A, que corresponde al valor dado de l, en la escala vertical. Dibujamos una perpendicular en este punto y marcamos un punto donde las dos perpendiculares se cortan. Cuando hayamos realizado la misma operación para cada uno de los valores de l y de A, tendremos la serie de puntos indicada en la Figura 1. Una línea curva continua se dibuja entonces, uniendo los diferentes puntos. Esta línea se denomina gráfica del área del cuadrado. Muestra cómo el área aumenta a medida que lo hace la longitud del lado.

Podemos usar la gráfica para determinar el área de cuadrados con lados no dados en la tabla de valores. Por ejemplo, si $l = 5{,}5$, elevamos una perpendicular en este punto, extendiéndola hasta que corte la gráfica. Desde el punto de corte, trazamos una perpendicular al eje vertical. El punto donde ésta corta el eje vertical representará el área, que es alrededor de 30. El valor correcto es 30,25.

Si conocemos el área de un cuadrado, podemos obtener la longitud aproximada del lado por medio de nuestra gráfica. Supongamos que el área es 75. Desde un punto correspondiente al número 75 en el eje vertical trazamos una línea paralela al eje horizontal. Desde el punto donde esta línea corta la curva, trazamos una perpendicular al eje horizontal. Esta línea cortará el eje horizontal aproximadamente en 8,7, la longitud del lado del cuadrado. Aquí, el valor correcto, expresado con cinco decimales, es 8,66025.

Hay muchas formas diferentes de gráficas. Las más comunes son: gráfica de barra, gráfi-



2. Gráficos de barra que muestran las poblaciones de cuatro países a mediados del siglo XX.

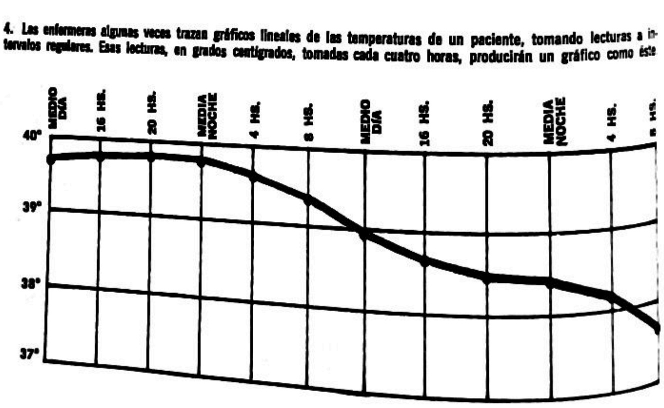
3. El gráfico circular indica cómo se gasta cada parte de un dólar destinado al departamento de una tienda

ca circular y gráfica lineal. Si, por ejemplo, quisiéramos comparar las poblaciones de China, India, la Unión Soviética y los Estados Unidos, dibujaríamos cuatro barras (Figura 2). Las longitudes de las barras serían proporcionales a las poblaciones en cuestión.

Si deseáramos comparar partes de una cantidad total, usaríamos un gráfico circular. Serviría para indicar cómo cada parte de un dólar recibido por ventas en una sección de una tienda es gastado por la tienda (Figura 3). Cada sector del círculo, comparado con el círculo completo, mostraría la proporción dada a un servicio particular.

Cuando una cantidad es continuamente variable, usaremos un gráfico lineal. En un hospital, la enfermera realiza un gráfico de la temperatura de cada paciente (Figura 4). Toma la temperatura cada cuatro horas, digamos, sitúa la temperatura leída en un punto dado, y luego une este punto con el punto precedente por medio de una línea recta. El médico que consulta

4. Las enfermeras algunas veces trazan gráficos lineales de las temperaturas de un paciente, tomando lecturas a intervalos regulares. Esas lecturas, en grados centígrados, tomadas cada cuatro horas, producirán un gráfico como éste



el diagrama puede ver de un vistazo cómo está cambiando la temperatura del paciente.

Resumiendo, pues, podemos mostrar cómo se relacionan las cantidades por medio de una fórmula, una tabla de valores o un gráfico. La fórmula es lo básico. Si se construye una tabla de valores, un científico trata de encontrar la fórmula que expresará la relación en cuestión. Si un ingeniero dibuja un gráfico que muestra la longitud de un cable de acero bajo una tensión creciente, determina una fórmula algebraica para sintetizar sus descubrimientos.

Las fórmulas son extremadamente importantes en muchos ramos de la ciencia pura y aplicada. Por ejemplo, para indicar la velocidad de un cuerpo que tiene movimiento rectilíneo uniforme –o sea, que se mueve en línea recta a velocidad constante– el físico usa la fórmula $v = \frac{d}{t}$, donde v es la velocidad promedio del cuerpo; d, el espacio o distancia recorridos y t, el tiempo requerido para recorrer esa distancia. Podemos aplicar esta fórmula a casos determinados sustituyendo valores apropiados para v, d o t. Por ejemplo, si un auto tarda 5 horas en recorrer 400 kilómetros, según indica el odómetro, podríamos encontrar la velocidad promedio sustituyendo 5 por t y 400 por d. La velocidad promedio sería, entonces, $\frac{400}{5}$ kilómetros, o sea 80 kilómetros por hora.

El químico tiene a menudo ocasión de usar las leyes de Charles y Gay-Lussac, que expresan que si la presión y la masa de un gas son constantes, el volumen es proporcional a la temperatura absoluta. La temperatura absoluta se basa en la temperatura más baja posible: "cero absoluto". El cero absoluto es igual a $-273{,}16°$ C. Esta ley puede expresarse muy brevemente por la fórmula $\frac{V_1}{T_1} = \frac{V_2}{T_2}$, en la que V_1 es el volumen de un gas a la temperatura T_1, mientras que V_2 es el volumen de un gas a la temperatura T_2.

Quizá la fórmula más famosa de todas es la ecuación de Einstein, $E = mc^2$. En esta ecuación, E representa la cantidad de energía, m la cantidad de masa, y c la velocidad de la luz (que es de aproximadamente 300.000 kilómetros por segundo), medidas en unidades apropiadas. Con esta fórmula, el gran físico del siglo XX Albert Einstein indicó la cantidad de energía que aparece cuando la materia se trasforma en energía. El uso de la energía atómica no hubiera sido posible sin esta fórmula.

CÓMO SE EMPLEAN LAS ECUACIONES

La ecuación representa un papel importantísimo en álgebra. Puede considerarse como una balanza, con iguales valores numéricos en cada lado del signo "igual" (=). Para mostrar cómo se aplica una ecuación, consideremos la fórmula del perímetro (el límite externo) de un rectángulo: $P = 2l + 2a$, donde P es el perímetro, l es el largo y a es el ancho. Supongamos que tenemos 100 metros de alambre para realizar un cercamiento rectangular, que debe ser de 20 metros de ancho. Deseamos averiguar la longitud de este cercado. Hemos visto que la fórmula para el perímetro es $P = 2l + 2a$. Sabemos que el perímetro es 100 (los 100 metros de alambre a nuestra disposición). También sabemos que el ancho debe ser 20 metros. Sustituyendo 100 por P y 20 por a en la fórmula, tenemos $100 = 2l + 2 \times 20 = 2l + 40$. En la ecuación $100 = 2l + 40$, el 100 a un lado del signo igual es exactamente igual a $2l + 40$ en el otro lado. Para averiguar cuánto es l, restamos 40 a cada lado de la ecuación. Eso da $60 = 2l$; por tanto, l es 30. La longitud del cerco es 30 metros.

Al resolver la ecuación, usamos la regla: "Si se realiza la misma operación en cada miembro de una igualdad, los resultados son iguales". Si sumamos la misma cantidad de cada lado de la ecuación o restamos la misma cantidad, la igualdad se mantendrá. También se mantendrá si multiplicamos o dividimos cada lado de la ecuación por la misma cantidad. Por supuesto, si multiplicamos cada lado de la ecuación por cero, el resultado sería $0 = 0$, que no nos llevaría a ninguna parte porque la división por cero no está permitida.

Las ecuaciones pueden usarse para resolver problemas en los que no está comprendida ninguna fórmula, pero en los que se dan ciertos datos. Aquí hay un problema sencillo: "Un hombre tiene 6 veces la edad de su hijo. Dentro de 20 años, el padre tendrá sólo 2 veces la edad de su hijo. ¿Cuántos años tienen el padre y el hijo actualmente?

De acuerdo con estos datos podemos escribir una ecuación y resolver el problema. Primero, representemos con x la edad del hijo. Como el padre tiene 6 veces la edad de su hijo, su edad puede escribirse 6 veces x, o sea, $6x$. En 20 años –o sea, cuando se agregan 20 años a la edad del hijo– el hijo tendrá $x + 20$ años. En 20 años, la edad del padre será $6x + 20$ años. En ese momento, la edad del padre será dos veces la del hijo, una relación que puede expresarse por la ecuación: $6x + 20 = 2(x + 20)$. Aplicando la ley distributiva al lado derecho de la ecuación, tenemos $6x + 20 = 2x + 40$. Restamos $2x + 20$ de cada lado de la ecuación y obtenemos $4x = 20$. Si $4x = 20$, $x = 5$. La edad del hijo actualmente es, entonces, 5. Como la edad del padre es actualmente 6 veces la del hijo, o sea $6x$, el padre tiene entonces 30 años.

No todos los problemas son tan sencillos como éste, en el que la incógnita es x. La ecuación puede implicar no sólo una cantidad incógnita, x, sino también potencias más elevadas de x. Si x^2 es la potencia más elevada que aparece en una ecuación, ésta se denomina una *ecuación cuadrática*. $x^2 + 6 = 5x$ es un ejemplo de dicha ecuación.

En otras ecuaciones, la potencia más alta de x puede ser x^3 o x^4, o tener aun potencias más altas.

IDENTIDADES

Cuando una ecuación es verdadera para todos los valores sustituidos a las variables implicadas, se denomina *identidad*. Un ejemplo muy común de identidad es $(a + b)^2 = a^2 + 2ab + b^2$. Como destacamos antes, esta ecuación se mantiene verdadera, no importa qué valores asignemos a a y b. Puede utilizarse como una ayuda en aritmética mental. Para elevar 22 al cuadrado, podemos pensar este número como $(20 + 2)^2$, sustituyendo 20 por a y 2 por b en la identidad de arriba. Mentalmente, elevamos 20 al cuadrado, lo que da 400; luego duplicamos 20×2, que da 80; finalmente, elevamos 2 al cuadrado, que es igual a 4; $400 + 80 + 4 = 484$. El resultado es en consecuencia, 484.

Otra identidad es $(a - b)^2 = a^2 - 2ab + b^2$. Esto se puede verificar realizando la multiplicación $(a - b)\ (a - b)$.

Tendríamos:

$$\begin{array}{r} a - b \\ a - b \\ \hline -\ ab + b^2 \\ a^2 - ab\quad\ \ \\ \hline a^2 - 2ab + b^2 \end{array}$$

Aún otra identidad es $(a + b)(a - b) = a^2 - b^2$. Esto también es útil en ciertos problemas de aritmética mental. Si deseamos multiplicar 34 por 26 mentalmente, podemos cambiar el problema a $(30 + 4) \times (30 - 4)$. Resolviendo esto de acuerdo con la identidad $(a + b)(a - b) = a^2 - b^2$, y tenemos: $900 - 16 = 884$. Otras identidades bien conocidas son:

$$(a + b)^3 = (a^3 + 3a^2b + 3ab^2 + b^3)$$
$$(a - b)^3 = (a^3 - 3a^2b + 3ab^2 - b^3)$$
$$(a^3 - b^3) = (a - b)(a^2 + ab + b^2)$$

EXPONENTES

Los exponentes simplifican la escritura de expresiones algebraicas. Así *aaabbbcc*, que es realmente una multiplicación sucesiva (a por a por a por b por b por b por c por c), puede escribirse $a^3b^3c^2$. Los matemáticos han obtenido una serie de reglas para combinar exponentes. Las reglas se expresan en términos generales: a^n representa la base a elevada a la *enésima potencia*; a^m, la misma base elevada a la *emésima potencia*.

(1) $a^na^m = a^{n+m}$. Al multiplicar potencias, sumamos los exponentes de bases iguales. Así $2^2 \times 2^3 = 2^{2+3} = 2^5 = 32$.

(2) $a^n \div a^m = a^{n-m}$. Al dividir potencias, restamos los exponentes de bases iguales. Eso significa que $2^5 \div 2^2 = 2^{5-2} = 2^3 = 8$.

(3) $(a^n)^m = a^{nm}$. Para elevar una potencia dada a otra potencia, multiplicamos los dos exponentes. Por ejemplo: $(2^2)^3 = 2^{2\times3} = 2^6 = 64$.

(4) $(ab)^n = a^nb^n$. Cuando se eleva un producto a una potencia, cada miembro del producto se eleva a esa potencia. Así $(4 \times 2)^2 = 4^2 \times 2^2 = 16 \times 4 = 64$.

(5) $\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}$. Cuando un cociente se eleva a una potencia dada, cada miembro del cociente debe elevarse a esa potencia.

$$\left(\frac{2}{3}\right)^3 = \frac{2^3}{3^3} = \frac{8}{27}.$$

Debe notarse aquí que una base con el exponente cero es equivalente a 1. Así $10^0 = 1$; $3^0 = 1$; $1^0 = 1$. Las reglas antedichas para exponentes se aplican a exponentes cero. Por ejemplo, $a^na^0 = a^{n+0} = a^n$; $5^3 \times 5^0 = 5^{3+0} = 5^3 = 125$.

Una base con un exponente negativo es igual a la recíproca de la base $\left(\frac{1}{\text{base}}\right)$ con el correspondiente exponente positivo. Así $2^{-2} = \frac{1}{2^2}$. Los exponentes negativos siguen la reglas para exponentes. Así $a^{-5}a^3 = a^{-5+3} = a^{-2} = \frac{1}{a^2}$;

$$10^{-7} \times 10^5 = 10^{-7+5} = 10^{-2} = \frac{1}{10^2} = \frac{1}{100}.$$

Los exponentes también pueden aparecer en forma de fracciones. Así tenemos $a^{1/2}$, $a^{1/3}$, $a^{1/4}$, y etc. $a^{1/2}$ significa la raíz cuadrada de a ($\sqrt{a}$); $a^{1/3}$ significa la raíz cúbica de a ($\sqrt[3]{a}$); $a^{1/4}$ significa la raíz cuarta de a ($\sqrt[4]{a}$). El numerador en exponentes fraccionarios no necesita ser 1. Frecuentemente tratamos con exponentes tales como $\frac{2}{3}$ y $\frac{3}{5}$. En tales casos, el numerador representa una potencia de la base, y el denominador, la raíz de una base. $10^{2/3}$, por ejemplo, es igual a $\sqrt[3]{10^2}$.

Todos los exponentes fraccionarios, sea o no 1 el numerador, siguen la regla para exponentes. Por ejemplo, $10^2 \times 10^{2/3} = 10^{2+2/3} = 10^{8/3} = \sqrt[3]{10^8}$.

EXPRESANDO NÚMEROS MUY GRANDES O MUY PEQUEÑOS

Los exponentes proporcionan un modo conveniente de escribir números muy grandes o muy pequeños. Sabemos que 1.000.000 es 10^6, donde el exponente 6 representa el número de ceros después de 1. Podríamos indicar 5.000.000 como $5 \times 1.000.000$ ó 5×10^6. Para escribir 5.270.000 multiplicaríamos 1.000.000 ó 10^6 por 5,27. El número se escribiría como $5{,}27 \times 10^6$. En otras palabras, podemos expresar un número grande como el producto de dos números: el primero, un número entre 1 y 10; el segundo, una potencia de 10.

El número 5.270.000 no es demasiado formidable y podemos entenderlo bastante fácilmente. Pero consideremos los problemas que podrían surgir si en nuestros cálculos tuviéramos que emplear un número tal como éste: 602.000.000.000.000.000.000.000. Representa el número de moléculas en 18 gramos de agua, y se denomina número de Avogadro, por el científico italiano Amedeo Avogadro, que a principios del siglo XIX determinó este valor. Se emplea en una gran cantidad de cálculos científicos, pero prácticamente nunca en la forma en que lo hemos dado. En lugar de ello, se escribe como $6{,}02 \times 10^{23}$.

El matemático norteamericano del siglo XX Edward Kasner inventó un nuevo sistema de indicar números extremadamente grandes. Acuñó la palabra "gúgol" para expresar el número 10^{100}, que sería equivalente a 1 seguido de 100 ceros. Inventó otro término, el "gúgolplex", para representar $10^{100^{100}}$, o la cifra 1 seguida por un gúgol de ceros elevado a un gúgol de ceros: o sea, 10.000 ceros.

Los exponentes pueden usarse con igual eficacia para expresar números muy pequeños, pues un exponente negativo indica cuántas veces la fracción $\frac{1}{\text{base}}$ se repite según se multiplica por sí misma; $10^{-3} = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} = 0{,}001$. Observamos que el exponente 3 en 10^{-3} representa el número de dígitos después de la coma decimal en 0,001. 0,005 podría escribirse $5 \times 0{,}001$ ó sea 5×10^{-3}. Consideremos ahora un número mucho más pequeño. La longitud de onda de la luz roja es 0,00000077 metros. Podemos escribir este número como $7{,}7 \times 0{,}0000001$ metros, ó $7{,}7 \times 10^{-7}$ metros.

La escritura de un número como el producto de (1) un número entre 1 y 10 y (2) una potencia de 10 se denomina *notación científica*. Es ampliamente utilizada por los científicos e ingenieros.

CÁLCULOS CON LOGARITMOS

Los exponentes también se han utilizado para simplificar los cálculos aritméticos. Supongamos que representamos números como potencias de 2. Sabemos que $2^{-2} = \frac{1}{2^2} = \frac{1}{4}$; 2^{-1}

$= \frac{1}{2^1} = \frac{1}{2}; 2^0 = 1;\ \ 2^1 = 2;\ \ 2^2 = 4;\ \ 2^3 = 8;$

$2^4 = 16$; etc., expresados como potencias de 2;

por tanto, $\frac{1}{4}$, ó 0,25, es 2^{-2}; $\frac{1}{2}$ ó 0,5, es 2^{-1}; 1 es

2^0; 2 es 2^1; 4 es 2^2; 8 es 2^3; 16 es 2^4. Hagamos ahora una tabla asentando: (1) ciertos números; (2) esos números expresados como potencias de 2; (3) los exponentes en cuestión.

Consideremos el problema $0{,}25 \times 1.024$. La tabla muestra que $0{,}25 = 2^{-2}$ y que $1.024 = 2^{10}$. El problema entonces se convierte en $2^{-2} \times 2^{10}$. Aplicando la primera ley de los exponentes dada más arriba, tenemos $2^{-2} \times 2^{10} = 2^{-2+10} = 2^8$. Consultando la tabla, hallamos que $2^8 = 256$. 256, entonces, es el resultado del problema $0{,}25 \times 1.024$. Hemos convertido un problema de multiplicación en un problema de suma: una operación más sencilla.

Tomemos otro problema: $1.024 \div 32$. Mirando la tabla, vemos que 1.024 es 2^{10} y que 32 es 2^5. Aplicando la segunda ley de los exponentes, tenemos $2^{10} \div 2^5 = 2^{10-5} = 2^5$. Ahora consultamos la tabla y encontramos que 2^5 es igual a 32. Éste es el resultado de $1.024 \div 32$. En este caso, hemos convertido un problema de división en un sencillo problema de resta.

Nuestro próximo problema es elevar 4 a la quinta potencia. En otras palabras, queremos saber qué sería 4^5. La tabla nos muestra que 4 es 2^2. Por la tercera ley de los exponentes, sabemos que $(2^2)^5 = 2^{2 \times 5} = 2^{10}$. De acuerdo con la tabla, 2^{10} es 1.024, que es el resultado. Hemos resuelto nuestro problema por una multiplicación sencilla en vez de multiplicar $4 \times 4 \times 4 \times 4 \times 4$.

Supongamos que queremos obtener la raíz cuadrada de 1.024. De acuerdo con la tabla, 1.024 es 2^{10}. Para obtener la raíz cuadrada de una potencia dada dividimos el exponente que indica esa potencia por 2. Por tanto, la raíz cuadrada de $2^{10} = 2^{10 \div 2} = 2^5$. Dividir el exponente que indica la potencia por 2 está realmente de acuerdo con la tercera ley de los exponentes. La raíz cuadrada de un número, como hemos visto, equivale al mismo número con el exponente $\frac{1}{2}$. La raíz cuadrada de 2^{10}, entonces, puede expresarse como $(2^{10})^{1/2}$. Recordemos que multiplicar un número por $\frac{1}{2}$ es lo mismo que dividirlo por 2. La tabla muestra que $2^5 = 32$; luego, 32 es la raíz cuadrada de 1.024.

Cuando un número se expresa como una potencia de una base dada, en este caso la base dos, llamamos al exponente que indica la potencia el *logaritmo* del número. Los exponentes en la tercera columna de la tabla son los logaritmos, en base dos, de los números de la primera columna. —2 es el logaritmo de 0,25 cuando la base es dos. Un matemático escribiría: $\log_2 0{,}25 = -2$. También, si la base es dos, el logaritmo de 4 es 2; el de 64 es 6. Para multiplicar números, sumamos sus logaritmos. Para dividir números, restamos sus logaritmos. Para elevar un número a una potencia dada, multiplicamos el logaritmo del número por la potencia en cuestión. Para obtener la raíz de un número, dividimos el logaritmo del número por la raíz deseada. Después que hemos sumado, restado, multiplicado o dividido en esta forma, hallamos el número que corresponde al logaritmo resultante.

Todos los logaritmos que dimos antes eran de base dos. La mayoría de las tablas de logaritmos se da en base diez. Preparemos ahora otra tabla, dando: (1) una serie de números; (2) los números expresados como potencias de 10; (3) los logaritmos de los números, o sea, los exponentes cuando los números se expresan como potencias de 10.

NÚMERO	NÚMERO EXPRESADO COMO POTENCIA DE 2	EXPONENTE EN LA COLUMNA ANTERIOR
.25	2^{-2}	−2
.5	2^{-1}	−1
1	2^0	0
2	2^1	1
4	2^2	2
8	2^3	3
16	2^4	4
32	2^5	5
64	2^6	6
128	2^7	7
256	2^8	8
512	2^9	9
1024	2^{10}	10

NÚMERO	NÚMERO EXPRESADO COMO POTENCIA DE 10	LOGARITMO DE BASE DIEZ (LOG_{10})
.0001	10^{-4}	−4
.001	10^{-3}	−3
.01	10^{-2}	−2
.1	10^{-1}	−1
1	10^0	0
10	10^1	1
100	10^2	2
1000	10^3	3
10000	10^4	4

Para resolver el problema $0{,}0001 \times 100$, consultamos la tabla y encontramos los logaritmos de 0,0001 y 100 (—4 y 2, respectivamente), sumamos los logaritmos ($-4 + 2 = -2$) y encontramos el número correspondiente al logaritmo —2. Este número, como vemos en la tabla, es 0,01. También podemos resolver problemas tales como $10.000 \div 0{,}0001$, 10^4 y $\sqrt{10.000}$.

Por supuesto, para ser útil, una tabla de logaritmos tendría que incluir otros números además de aquellos dados más arriba, o sea, por ejemplo, no sólo los logaritmos de 1 y 10, sino también los de 2, 3, 4, 5, 6, 7, 8 y 9. Sabemos que como $1 = 1^0$ y $10 = 10^1$, el logaritmo de 2 estaría entre 0 y 1. Los matemáticos han calculado que es 0,301. Esto significa que, expresado como una potencia de 10, el número 2 es $10^{0.301}$. La parte entera del logaritmo (0 en este caso) se denomina *característica*; la parte decimal (,301), *mantisa*.

Damos sólo tres cifras decimales para mayor facilidad. Los logaritmos se han calculado con más de veinte decimales. Dependiendo de la aproximación deseada, uno usaría una tabla de cuatro decimales, o una de cinco decimales, o una de siete decimales y así sucesivamente.

Los logaritmos de los demás números de 3 a 9 también se han calculado, como puede verse en la tabla al pie.

NÚMERO	LOGARITMO (BASE DIEZ)
1	0.
2	0.301
3	0.477
4	0.602
5	0.699
6	0.778
7	0.845
8	0.903
9	0.954
10	1.

Usando la tabla, multipliquemos 2 por 4. Sumamos 0,301, el logaritmo de 2, y 0,602, el logaritmo de 4, y obtenemos el logaritmo 0,903. Consultando la tabla, vemos que 0,903 es el logaritmo de 8. Un matemático diría que 8 es el antilogaritmo, o antilog., de 0,903. Un *antilogaritmo* es el número que corresponde a un logaritmo dado. 8, por tanto, es la respuesta al problema 2×4. Dividamos ahora 9 por 3. El logaritmo de 9, como vemos en la tabla, es 0,954. El logaritmo de 3 es 0,477. Restando 0,477 de 0,954, obtenemos 0,477. La tabla muestra que 0,477 es el logaritmo de 3. Por tanto, $9 \div 3 = 3$.

Nuestra tabla de números y logaritmos de base diez da sólo diez números. Los matemáticos han preparado tablas que hacen posible encontrar el logaritmo de cualquier número. Las tablas dan sólo las mantisas. Podemos determinar la característica en cada caso por inspección. Por ejemplo, el logaritmo del número 343 debe estar entre 2 y 3, pues $100 = 10^2$ y $1.000 = 10^3$. El logaritmo, entonces, debe ser 2 y una fracción. La característica debe ser 2. Si consultamos una tabla de cinco decimales con el propósito de encontrar la mantisa de 343, observamos que es igual a 0,53529. Uniendo ambas, la característica 2 y la mantisa 0,53529, tenemos el logaritmo 2,53529.

En los cálculos que comprenden problemas aritméticos, a menudo podemos ahorrar gran cantidad de tiempo consultando una tabla de logaritmos. Por supuesto que no usaríamos logaritmos para resolver 4×5 ó $72 \div 9$. Pero supongamos que tuviéramos que realizar las diversas operaciones de un problema como el que se expone a continuación:

$$\frac{-2{,}953 \times 5{,}913^5 \times \sqrt[5]{5{,}973}}{49{,}743 \times 0{,}35947^3}$$

Si se usaran métodos aritméticos, tal problema sería una tarea muy laboriosa. Podría realizarse en unos pocos minutos si empleáramos logaritmos.

Los logaritmos de base diez se denominan *logaritmos comunes*. Se han utilizado otras bases. En los llamados *logaritmos naturales*, la base es 2,71828 . . . , generalmente indicada por la letra *e*. Estos logaritmos naturales sirven ampliamente en varios tipos de análisis superior porque conducen a fórmulas comparativamente sencillas.

SECUENCIAS ALGEBRAICAS Y SERIES

Muchos sucesos parecen repetirse en un orden regular. El Sol "se levanta" todos los días. Los planetas giran en sus órbitas alrededor del Sol tan regularmente que los astrónomos pueden calcular su posición con años de adelanto. Se han analizado los sucesos periódicos por medio de sucesiones y series algebraicas. Una *secuencia* es una sucesión de números. Una *serie* es una suma de números en una secuencia. Los resultados de estos análisis se usan algunas veces para predecir sucesos futuros. Consideraremos brevemente aquí las secuencias aritmética y geométrica y la serie binómica.

La serie aritmética. En una secuencia aritmética, cada término, después del primero, se forma sumando una cantidad constante al término precedente. Un ejemplo de dicha secuencia es 1, 3, 5, 7, 9, 11, 13, 15, en la que 2 se suma a cada número subsiguiente de la secuencia. Si a representa el primer término; d, el número constantemente sumado, y n, el número total de términos, podemos representar la secuencia aritmética algebraicamente por esta fórmula:

$$a,\ (a+b),\ (a+2d)\ldots[a+(n-1)d]$$

Aquí $[a+(n-1)d]$ es el término *enésimo*.

Si sumamos n términos, dicha suma se denomina una serie y puede expresarse en la fórmula:

$$s(\text{suma})=\frac{n}{2}[2a+(n-1)d]$$

Apliquemos esta fórmula a la suma de los términos en la secuencia dada más arriba: 1, 3, 5, 7, 9, 11, 13, 15. Aquí hay ocho términos en total. El primer término es 1. La cantidad constantemente sumada es 2. Sustituyendo estos valores para n, a, y d:

$$s=\frac{8}{2}[2\times 1+(8-1)\times 2]$$

Si resolvemos el cálculo aritmético implicado, se encontrará que la suma de los ocho términos es 64. Se puede verificar sumando los ocho términos de la secuencia.

La serie aritmética es muy útil en varios tipos de cálculos. Sirve, entre otras cosas, para hallar el precio total de un artículo que se ha comprado en mensualidades. Supongamos que alguien adquiere un piano por $ 1.000, paga $ 400 al contado y se compromete a pagar los otros $ 600 en 20 cuotas mensuales de $ 30 cada una, más el interés de 6 por ciento por las cuotas no saldadas. Apliquemos la serie aritmética al problema para determinar el pago total de interés que se requerirá.

El 6 por ciento de interés significa "6 por ciento anual". El período de plazo es un mes, o $\frac{1}{12}$ de año. El primero de estos pagos es $\frac{1}{12}\times 0{,}06\times \$\ 600$ (saldo no pagado), o sea $ 3,00. Cada mes, el interés es menor que en el mes precedente, pues el saldo no pagado se reduce en $ 30. Se pagará $\frac{1}{12}\times 0{,}06\times \$\ 30$, o $ 0,15 menos de interés que el mes anterior. Los pagos de interés, por tanto, serían $ 3,00 (para el primer mes), $ 2,85 (para el segundo mes), $ 2,70 (para el tercer mes) y así sucesivamente hasta que se pagaran las 20 cuotas. Volviendo a la fórmula de la suma de una serie aritmética, vemos que n (el número de términos) es 20 en este caso; que a (el primer término) es $ 3,00; que d (el número sumado a los diferentes términos) es $-0{,}15$. Haciendo las sustituciones apropiadas en la fórmula, tenemos:

$$s=\frac{20}{2}\times[(2\times 3{,}00)+(20-1)\times -0{,}15]$$

El resultado, que representa el interés total pagado, es $ 31,50.

La secuencia geométrica. En una secuencia geométrica, cada término, después del primero, se forma multiplicando el anterior por una misma cantidad fija. Una secuencia geométrica típica es 1, 2, 4, 8, 16, en la que el término fijo usado como multiplicador es 2. Algebraicamente, cualquier secuencia geométrica puede representarse como:

$$a,\ ar,\ ar^2,\ldots ar^{n-1}$$

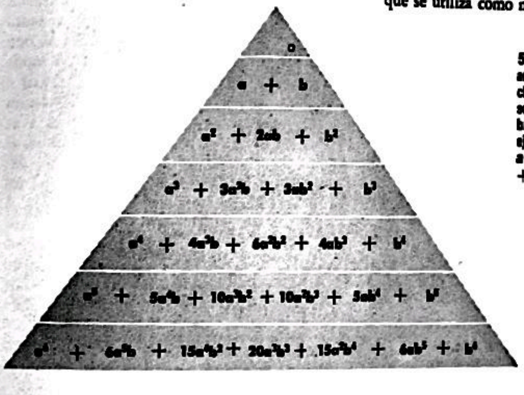
donde a es el primer término; r, el multiplicador constante, y n, el número de términos. La suma de n términos de una secuencia geométrica –una suma llamada serie geométrica– se expresa por la fórmula:

$$s=\frac{ar^n-a}{r-1}$$

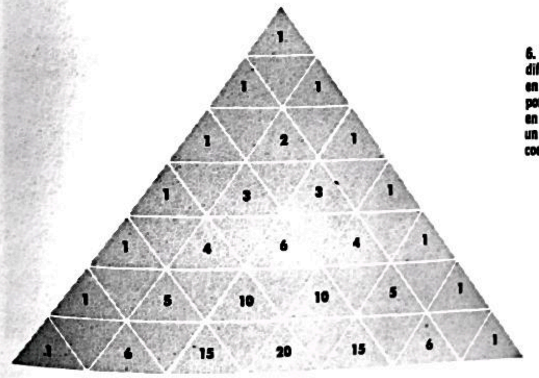
Aplicándola a la serie $1+2+4+8+16$, se encontrará que la suma es 31.

La serie geométrica tiene un papel importante en las matemáticas de las finanzas. Se usa, entre otras cosas, para calcular el interés compuesto. Supongamos que ponemos $ 100 en un banco y que el interés es 3 por ciento, compuesto anualmente. El interés se dice compuesto cuando se aplica, no al capital solamente, sino al capital más el interés no pagado, periódicamente sumado al capital. Al final del primer año, el capital más el interés se convertiría en $ 100 más $ 3,00 (que representa el interés), o un total de $ 103. Podríamos escribir esto como $ 100 × 1,03. Para el período siguiente –o sea, para el segundo año– recibiríamos interés de 3 por ciento sobre $ 103. Así la cantidad total al final del segundo año sería $ 103 + (0,03 × $ 103). Expresada en forma algo diferente, la cantidad en cuestión sería igual a $\$\ 100\times 1{,}03^2$.

Para repasar, entonces: al final del primer año, la cantidad total que tendríamos en el banco sería $ 100 × 1,03; al final del segundo año, tendríamos $\$\ 100\times 1{,}03^2$. Siguiendo el mismo procedimiento, tendríamos $\$\ 100\times 1{,}03^3$ al final del tercero; $\$\ 100\times 1{,}03^4$ al final del cuarto, etc. Esto representa la secuencia geométrica: $ 100, $ 103, $ 106,09, $ 109,27, $ 112,55, ... en la que 1,03 es el término fijo que se utiliza como multiplicador.



5. Una serie binómica leída de arriba abajo. Cada miembro de dicha serie, después del primero, se forma multiplicando el miembro precedente por $a+b$. Por ejemplo, $a+b$, multiplicado por $a+b$, es igual a $a^2+2ab+b^2$.



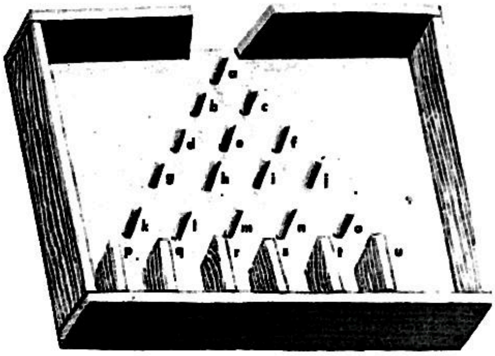
6. Aquí, los coeficientes de las diferentes potencias de a y b, en una serie binómica, se disponen en el mismo orden que en la figura 5. Un coeficiente es un factor. Por ejemplo, 2 es el coeficiente en $2ab$.

Al final de *n* períodos, la cantidad total de dinero en el banco sería \$ 100 × $1{,}03^n$. Para diez períodos de interés, la suma se convertiría en \$ 100 × $1{,}03^{10}$. Calculando por logaritmos, encontramos que $1{,}03^{10} = 1{,}344$. Y 1,344 × \$ 100 = 134,40; \$ 100, al 3 por ciento de interés compuesto anual, sería \$ 134,40 al término de diez años.

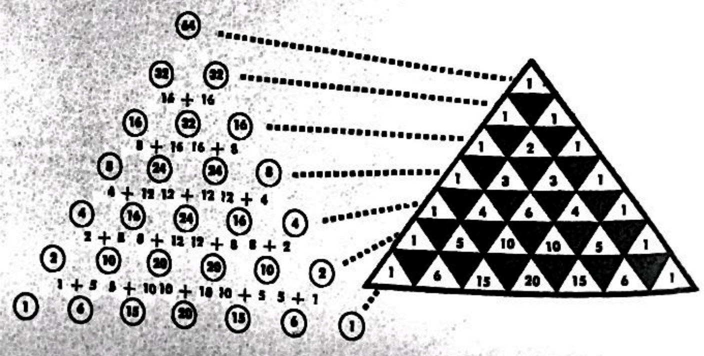
La serie binómica. Un binomio consiste en dos términos conectados por un signo más o menos. $a + b$, $2x + z$, $x^2 - y^2$ son todos binomios. La serie binómica es la suma de términos basada en el desarrollo de la potencia de un binomio. La serie para $(a + b)^1$ es $a + b$. Para $(a + b)^2$, obtenemos la serie multiplicando $a + b$ por $a + b$, lo cual da $a^2 + 2ab + b^2$. La serie para $(a + b)^3$ se obtiene multiplicando $a^2 + 2ab + b^2$ por $a + b$. Esto da $a^3 + 3a^2b + 3ab^2 + b^3$. Podemos disponer esta serie binómica y también aquéllas para $(a + b)^4$, $(a + b)^5$ y $(a + b)^6$ como en la Figura 5.

Supongamos ahora que, en vez de distintas potencias de *a* y *b*, damos sólo los coeficientes –números o letras usados como multiplicadores– y reproducimos el triángulo en esta forma modificada como muestra la Figura 6. Se notará que agregamos el número 1 en el vértice del triángulo.

7. Los tacos (a, b, c, d, e, f, etc.), en esta caja poco profunda, están en las posiciones exactas de los números de la figura 6. Las paredes que se extienden desde los tacos del fondo forman una serie de compartimientos en el fondo de la caja.



8. Si dejamos caer 64 discos, uno por uno en la caja que muestra la figura 7, realizarán su camino descendente hacia los tacos, como está indicado. Las relaciones de los discos que golpean los tacos en cada fila (derecha) reproducen el triángulo de la Figura 6.



Coloquemos ahora una serie de tacos en una caja poco profunda, en las posiciones exactas de los números del triángulo numérico, como se muestra en la Figura 7. Los tacos deben estar suficientemente alejados para que un pequeño disco pueda pasar entre ellos. Recortamos ahora una pequeña sección de un lado de la caja (que será el tope), como se indica en la figura, y también construimos una serie de paredes, desde cada taco del fondo, para formar una serie de compartimientos.

Mantenemos la caja en una posición inclinada, para que cuando un disco se deje caer a través de la abertura superior, recorra su camino a través del laberinto de tacos hacia uno de los compartimientos del fondo. Dejamos ahora caer 64 discos, uno por uno, en la caja a través de la ranura. Podemos esperar que la mitad de los discos que golpean en un taco caerá a la izquierda de éste y la otra mitad, a la derecha. Así, 32 discos caerían a la izquierda del taco *a* y 32, a la derecha del mismo. De los 32 que caen a la izquierda y golpean al taco *b*, 16 podrían caer a la izquierda de *b* y golpearían el taco *d*; 16 podrían caer a la derecha del taco *b* y golpearían el *e*. De los 32 discos que golpean el taco *c*, 16 golpearían el taco *e* y 16 golpearían el *f*. Eso significa que 32 discos en total se encontrarán con el taco *e*. Podemos indicar cómo caerían en su recorrido hacia los compartimientos del fondo por el diagrama en la Figura 8.

Como hemos visto, en la segunda fila, 32 discos golpearían el taco *b* y 32, el taco *c*. 32 y 32 están en la razón 1-1. En la tercera fila, 16 golpearían *d*, 32 golpearían *e* y 16 golpearían *f*. La razón entre 16, 32 y 16 es 1-2-1. Descendiendo por las filas, las razones de los discos que golpean los diferentes tacos serían 1-3-3-1 en la cuarta fila, 1-4-6-4-1, en la quinta y 1-5-10-10-5-1, en la sexta. Los discos en la séptima fila seguirían la distribución 1-6-15-20-15-6-1. Notemos que estas razones corresponden todas exactamente a los coeficientes de la serie binómica en la Figura 6.

En un experimento real, los discos no caerán exactamente como hemos indicado. Algunos compartimientos tendrán uno o dos más que el número previsto. Otros tendrán uno o más, menos. No obstante, en cada caso el resultado será aproximadamente aquel que fue previsto. Si el experimento se repite una y otra vez, en un gran número de pruebas, el número en cada compartimiento coincidirá más y más estrechamente con el número esperado. Así, los coeficientes de la serie binómica proveen un medio efectivo para calcular probabilidades cuando las posibilidades de que ocurra un acontecimiento son uniformes.

USADO EN ESTADISTICA Y PROBABILIDAD

Para determinar las probabilidades cuando las posibilidades no son uniformes se ha realizado otro análisis algebraico. La disciplina que trata dicho análisis se denomina *estadística*. Se usa en muchos campos, incluyendo seguros. Las compañías de seguros pueden, por ejemplo, predecir con bastante precisión los promedios de vida, y cuánto en premios e interés recaudarán durante el tiempo que vivirá una persona asegurada. A fin de trabajar con fruto, la compañía debe recaudar lo suficiente para poder pagar el seguro cuando el cliente muera. La probabilidad y el interés compuesto, así determinados por la estadística, son por ende las bases sobre las que se realiza el seguro. Es así cómo el álgebra, que comenzó examinando las relaciones de las operaciones aritméticas, se ha vuelto un intérprete de nuestra experiencia y una guía para el futuro.

Thomas R. Taylor, Photo Researchers

En la naturaleza, la ciencia y el arte pueden encontrarse ejemplos de geometría en variedad infinita. Arriba, izquierda: el caparazón de un nautilo con cámaras forma una espiral cuyas dimensiones crecen uniformemente. Arriba, derecha: la Gran Pirámide de Gizeh es un testimonio del conocimiento de la geometría en los pueblos antiguos.

GEOMETRÍA PLANA

Cuando pasamos de la aritmética y el álgebra a la geometría penetramos en un mundo de formas que se encuentran en el espacio: un conjunto de puntos, líneas, superficies y sólidos. Estudiamos las propiedades de estas formas y las relaciones que existen entre ellas. Aprendemos a medirlas. Al principio, la geometría se empleaba para resolver problemas determinados, pero, en el curso de su desarrollo, se convirtió en una materia completamente abstrusa; empero, a este abstruso ramo de las matemáticas, como veremos, se le puede dar con frecuencia empleo práctico.

Los comienzos de la geometría se remontan a la prehistoria. A medida que creció la población de una región dada, los lugares naturales disponibles para habitación se volvieron insuficientes; fue necesario construir refugios bastante grandes para alojar familias, y con la suficiente resistencia para soportar vientos, lluvia y tormentas. Para construir un refugio de tamaño adecuado, una persona debía comparar longitudes: así, el techo tenía que estar más alto que el extremo de la cabeza de la persona de mayor estatura.

Los antiguos babilonios fueron pioneros en este ramo de las matemáticas. La tierra comprendida entre los ríos Tigris y Éufrates –donde ese pueblo habitaba– era originalmente un pantano; se construyeron canales para drenar los marjales y para contener el desbordamiento de los ríos. A fin de construir canales fue necesario planificar el terreno; al hacerlo, los babilonios desarrollaron reglas para el cálculo de superficies, las que, si bien no eran exactas según los criterios modernos, bastaban para la construcción de canales.

En Egipto se gravaba a la gente que tenía granjas a lo largo de las márgenes del río Nilo, en función de sus posesiones. Durante la estación de las lluvias, el río desbordaba de sus riberas y borraba todos los hitos; en consecuencia, era necesario volver a medir la tierra, a fin de que todo propietario tuviese su porción legítima. Una vez que las inundaciones habían retrocedido, hombres especialmente entrenados –llamados tendedores de cuerda– establecían nuevos hitos: empleaban sogas anudadas a intervalos regulares, para poder medir las longitudes deseadas, y dividir la tierra en triángulos, rectángulos y trapecios.

Los egipcios crearon una variedad de reglas de uso práctico para medir las superficies de estas figuras, reglas que eran elementales y fáciles, pero, con frecuencia, inexactas. Hoy en día sabemos, por ejemplo, que la superficie de cualquier triángulo es igual a la mitad del producto de su altura por su base: erróneamente, los egipcios expresaron esta superficie como la mitad del producto de la base por un lado. Sin embargo, la mayor parte de los triángulos que emplearon en su tarea de relevamiento eran largos y estrechos (Figura 1), y en tales triángulos no hay mucha diferencia de longitud entre el lado y la altura. Por eso, el resultado de los cálculos hechos por ellos sirvió como base bastante buena para la asignación de tierras y la imposición de gravámenes a los terratenientes.

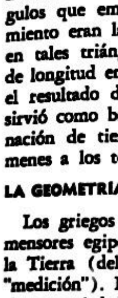
LA GEOMETRÍA SE TRASFORMA EN DISCIPLINA

Los griegos llamaron a los primitivos agrimensores egipcios geómetras, o medidores de la Tierra (del griego *ge:* "Tierra", y *metria:* "medición"). Los geómetras descubrieron muchas propiedades de los triángulos, cuadrados, rectángulos e, inclusive, círculos, las que constituyeron un cuerpo de conocimientos que los griegos denominaron geometría, o "estudio de la medición de la Tierra". En la actualidad, la geometría incluye mucho más que en su etapa inicial, pero todavía trata de tamaños, formas y posiciones de las cosas.

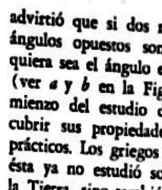
Los griegos hicieron importantes progresos en el campo de la geometría; no sólo corrigieron muchas de las reglas erróneas de los egipcios, sino que también estudiaron las diferentes figuras geométricas, con el objeto de determinar relaciones. Tales, matemático griego que vivió hace 2.500 años, descubrió que si se traza un diámetro cualquiera en un círculo, siempre lo biseca: es decir, corta el círculo en dos mitades (Figura 2). También advirtió que si dos rectas se entrecruzan, los ángulos opuestos son siempre iguales, cualquiera sea el ángulo en que aquéllas se crucen (ver *a* y *b* en la Figura 3). Esto fue el comienzo del estudio de las figuras para descubrir sus propiedades, y no con propósitos prácticos. Los griegos cambiaron la geometría; ésta ya no estudió solamente la medición de la Tierra, sino también las relaciones entre las diversas partes de las figuras que existen en el espacio; hoy damos a la geometría este significado.

Después de Tales, otros matemáticos griegos descubrieron y probaron propiedades de las figuras geométricas: las expusieron en enunciados llamados *teoremas*. Asimismo, idearon varios instrumentos para el trazado de figuras; conforme a la tradición, los únicos instrumentos admitidos en el estudio formal de la geometría eran un listón recto sin marcas, o regla, para trazar rectas, y un compás para dibujar círculos y trasportar mediciones (Figura 4).

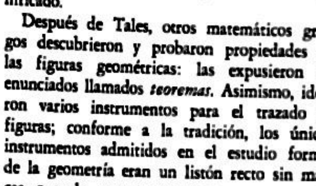
Los griegos propusieron varios problemas de construcción que debían resolverse con regla y compás solamente. Entre éstos figuran: (1) construir un cuadrado cuya superficie fuese exactamente igual a la de un círculo dado: lo que se llama cuadratura del círculo. (2) Construir la arista de un cubo cuyo volumen sea exactamente el doble del de un cubo dado: lo que se denomina duplicar el cubo. (3)



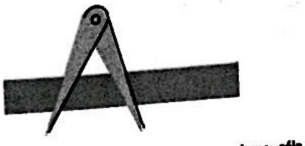
1. En este triángulo largo y estrecho, cada lado no es mucho más largo que la altura. Los egipcios usaban tales triángulos para reconocimiento topográfico.



2. Cada uno de los diámetros que se muestra en el dibujo corta el círculo en mitades.



3. Si se cruzan dos rectas, los ángulos opuestos siempre son iguales. En a, ángulo 1 = ángulo 2, y ángulo 3 = ángulo 4. En b, 5 = 6 y 7 = 8.



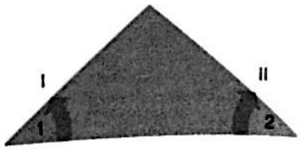
4. Conforme a la tradición, los antiguos griegos sólo usaban en geometría una regla sin marcas y un compás.

Construir un ángulo exactamente igual a un tercio de un ángulo dado: lo que se denomina trisecar al ángulo. Durante más de veintidós siglos los matemáticos intentaron, infructuosamente, resolver estos problemas. Finalmente, en el siglo diecinueve, se probó que era imposible cuadrar el círculo, duplicar el cubo o trisecar un ángulo si sólo se recurre a regla y compás; empero, es factible hacer estas tres construcciones con instrumentos especialmente diseñados; estas construcciones pertenecen al dominio de la geometría superior.

EUCLIDES: ORGANIZADOR DE LA GEOMETRÍA

Alrededor del siglo IV a.C. se había realizado una vasta recopilación de hechos concretos concernientes a las figuras geométricas, pero en su mayor parte no estaban relacionados. Existían muchos teoremas sobre triángulos y círculos, algunos acerca de superficie y figuras semejantes, pero no había una disposición ordenada. El ilustrado matemático griego Euclides, que enseñó en el Museo de Alejandría, Egipto, alrededor del 300 a.C., fue el primero que aplicó un desarrollo lógico al conocimiento matemático de su tiempo y lo presentó en sus *Elementos de geometría.*

Euclides verificó que no es posible probar cada cosa que decimos, y que debemos dar algunas por supuestas. Partió de la presunción de que todos saben, y emplean correctamente, palabras tales como "entre", "sobre", "punto" y "línea", por lo que no es necesario definirlas; empleó estos términos no definidos para dar las definiciones de varias figuras. Así, definió un círculo como "una línea curva cerrada cuyos puntos equidistan de uno fijo llamado centro". Nuevamente, Euclides observó que no se pueden probar ciertas proposiciones sobre relaciones entre figuras geométricas; un ejemplo sería: "Entre dos puntos sólo puede trazarse una recta." Euclides llamó a tales enunciados nociones comunes: pero,



5. Si dos lados de un triángulo son iguales, los ángulos opuestos a ellos también lo son: si I = II, 1 = 2.

actualmente, nosotros las llamamos *postulados.*

Euclides empleó términos no definidos, definiciones y postulados para probar *teoremas* sobre figuras geométricas. Un teorema es un enunciado que describe ciertas propiedades de una figura y deduce a partir de ellas que cierta otra propiedad tiene que ser verdadera. Un teorema típico es: "Si dos lados de un triángulo son iguales, los ángulos opuestos a ellos también deben serlo" (Figura 5); el teorema establece la premisa de que (1) existe un triángulo y que (2) dos lados del mismo son iguales; luego, extrae como conclusión que dos de los ángulos del triángulo son iguales. Una vez probado un teorema, puede emplearse para probar otros.

Euclides estructuró una cadena lógica de teoremas que ordenaron lo que había sido un caos de hechos más o menos desconectados. Amén de organizar un vasto cuerpo de conocimientos con respecto a las figuras geométricas, introdujo un método de tratamiento que llegó a ser modelo para el desarrollo de otros ramos de la matemática y la ciencia pura; este método es tan válido hoy como entonces.

DOS DIMENSIONES

El primer ramo de la geometría que hemos de considerar es la geometría plana; el estudio de puntos, líneas y figuras que se hallan en planos. ¿Qué es, exactamente, lo que queremos decir con estas palabras?

Un *punto* es el elemento más simple en geometría: carece de longitud, ancho o espesor, lo que es otro modo de expresar que no tiene ninguna dimensión. Podemos representar un punto por una marca hecha con lápiz o con un trozo de tiza. Tal marca no es un punto geométrico, sino físico, puesto que tiene longitud, ancho y espesor, no importa cuán pequeñas puedan ser estas dimensiones. En las construcciones geométricas tenemos que usar puntos físicos –tales como marcas de lápiz– para representar los geométricos, dado que nos sería imposible marcar en el papel un punto sin dimensiones.

Si hay dos puntos diferentes, la distancia más corta entre ellos es una *recta.* Este segmento de recta sólo tiene una dimensión, llamada *longitud:* carece de ancho y espesor. La recta que tracemos con lápiz sobre un papel sí los tiene, por lo que, cuando dibujamos una línea al construir una figura geométrica, estamos de nuevo representando físicamente un elemento geométrico.

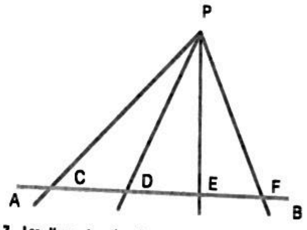
Si estuviésemos confinados a un mundo de tan sólo una dimensión –tal como la longitud– lo pasaríamos bastante aburridos: seríamos puntos de una recta, y podríamos sólo desplazarnos hacia adelante o hacia atrás, tropezando siempre con puntos delante o detrás de nosotros. En la Figura 6, los puntos *A* y *B* y el segmento *CD* son todos parte de la recta que se muestra en la figura, y siempre deben estar dentro de ella.

Supongamos ahora que hemos elegido un punto *P* exterior a la recta; las líneas trazadas desde dicho punto *P* a la recta original crean una serie de figuras que se hallan en un plano. Un *plano* es una superficie que tiene dos dimensiones: largo y ancho (Figura 7); la superficie de una mesa es un plano. Una continuación de la superficie representaría parte de ese mismo plano. Si fuésemos puntos de un mundo bidimensional, podríamos movernos libremente en cualquier dirección, excepto hacia afuera del plano. Nuestro mundo tendría otros puntos como nosotros mismos, y también líneas. Asimismo, habría gran variedad de figuras –constituidas por combinaciones de puntos y rectas– tales como triángulos, cuadrados, círculos, etcétera.

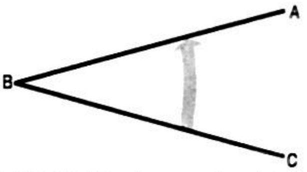
ÁNGULOS EN GEOMETRÍA PLANA

Se da el nombre de *rayo* a la parte de una recta que comienza en un punto dado. Una figura plana formada por dos rayos que tienen un mismo punto de origen se denomina *ángulo.* En la Figura 8, *AB* y *BC* son dos rayos con el mismo origen, *B*. El ángulo formado por ambos rayos es el *ABC*. Ustedes observarán que la letra *B* –que representa el origen– se intercala entre la letra *A* –de uno de los rayos– y la *C*, del otro; así se indican siempre los ángulos.

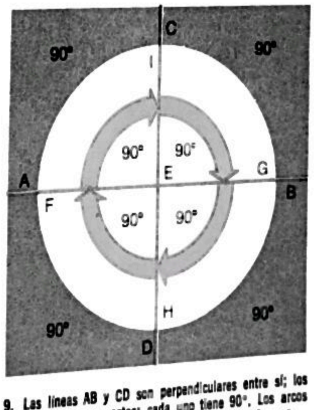
6. Los puntos A y B y la longitud CD son partes de la recta de arriba, y deben quedar siempre dentro de ella.



7. Las líneas trazadas desde P a la recta AB crean varias figuras (tales como PCD), que se hallan en un plano.



8. Arriba, AB y BC son dos rayos que tienen el mismo origen B; forman el ángulo ABC.



9. Las líneas AB y CD son perpendiculares entre sí; los ángulos en E son rectos; cada uno tiene 90°. Los arcos cortados en el círculo también tienen 90° cada uno.

Si se cortan dos rectas de tal modo que todos los ángulos formados sean iguales, se dice que las rectas son *perpendiculares*, y a los ángulos se los llama *ángulos rectos*; en la Figura 9, *AB* es perpendicular a *CD*, y los cuatro ángulos –*AEC*, *BEC*, *AED* y *BED*– son iguales. Si trazamos un círculo en torno al punto *E*, su longitud –llamada *circunferencia*– puede dividirse en 360 unidades, llamadas grados, que se escriben con el símbolo °. Las partes del círculo designadas *IG*, *GH*, *HF* y *FI* se llaman *arcos*, cada uno de los cuales tiene 90°, puesto que la circunferencia del círculo está dividida en cuatro partes iguales por las líneas *AB* y *CD*. El ángulo central del círculo tiene el mismo número de grados que el arco que corta en el círculo, de donde resulta que cada uno de los cuatro ángulos mencionados anteriormente tiene 90°; en otras palabras, un ángulo recto mide 90°.

Si un ángulo es menor que un recto –es decir, menor de 90°– se denomina *agudo*, y es *obtuso* si es mayor que un ángulo recto: es decir, de más de 90°. Cuando el ángulo obtuso es tan abierto que sus lados forman una recta, se lo llama *ángulo llano*, y tiene 180°. Un ángulo más grande que uno llano es un *ángulo convexo:* por supuesto, debe tener más de 180°. La Figura 10 muestra estas diferentes clases de ángulos, que pueden medirse con el instrumento denominado trasportador, el cual consiste en un semicírculo dividido en 180 partes, cada una de las cuales representa un grado de ángulo central (Figura 11). Como veremos, los ángulos tienen un papel preponderante en el estudio de la geometría.

EL ESTUDIO DE LOS TRIÁNGULOS

Una parte determinada de una recta –la distancia entre dos puntos particulares– se llama *segmento de recta*. Cuando tres de ellas unen tres puntos de un plano no alineados, forman un *triángulo*. La Figura 12 muestra las tres clases diferentes de triángulos que pueden formarse. Literalmente, existen miles de teoremas sobre los lados, ángulos y rectas de triángulos.

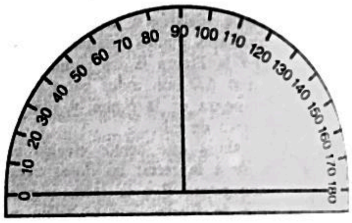
Uno de los primeros teoremas que se probaron en geometría plana es éste: "Si se dan tres longitudes definidas tales que la suma de dos cualesquiera de ellas sea mayor que la tercera, es posible utilizar dichas longitudes para construir un triángulo que tenga forma y tamaño definidos" (Figura 13). Como la forma nunca varía, una construcción hecha en forma de triángulo será rígida y no "cederá". A causa de esta propiedad, en todos los diseños de puentes y edificios se usan triángulos entrelazados, a fin de evitar que una estructura se desplome. Una figura de cuatro lados de longitud definida cada uno podría asumir muchas formas diferentes, tal como se muestra en la Figura 14; una construcción hecha según esta forma podría derrumbarse, por lo que no puede empleársela para construcciones rígidas, a menos que esté reforzada con una diagonal, que convierte a una figura cuadrilátera en dos triángulos (Figura 15), cada uno de los cuales es rígido.

El teorema más famoso –y, quizás, el más importante– de la geometría plana se refiere a un triángulo rectángulo (aquel que tiene un ángulo recto): se lo llama teorema de Pitágoras por su descubridor, el filósofo griego Pitágoras, que vivió en el siglo VI a.C. Este teorema afirma que "La suma de los cuadrados de dos lados de un triángulo rectángulo es igual al cuadrado de la hipotenusa (lado opuesto al ángulo recto)". En el triángulo rectángulo de la Figura 16, los lados tienen 3, 4 y 5 unidades de longitud: el lado de 5 unidades es la hipotenusa. Trazamos los tres grandes cuadrados como se indica: el primero con un lado (*AB*) de 3 unidades; el segundo con un lado (*BC*) de 4; el tercero con un lado (*AC*) de 5. Según el teorema de Pitágoras, $(AB)^2 + (BC)^2 = (AC)^2$. En este caso $3^2 + 4^2 = 5^2$, ó $9 + 16 = 25$. Podemos verificarlo contando los cuadrados pequeños –cada uno con lados de una unidad de largo– en los tres cuadrados grandes; hay 9 cuadrados pequeños en el *ABED*, 16 en el *BCGF*, y 25 en el *ACIH*.

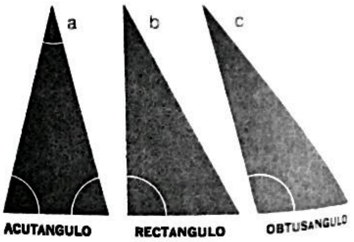
La Figura 17 muestra cómo pueden cortarse los dos cuadrados más pequeños de los lados del triángulo rectángulo, de modo que formen el cuadrado de la hipotenusa; ésta es otra confirmación del teorema de Pitágoras.

Se desprende de este teorema que si un triángulo tiene lados tales que la suma de los cuadrados de los dos menores sea igual al cuadrado del mayor, el ángulo opuesto a este último es recto. Este teorema tiene varias apli-

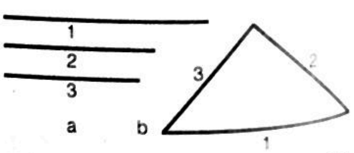
10. Los diferentes tipos de ángulo, de agudo a convexo.



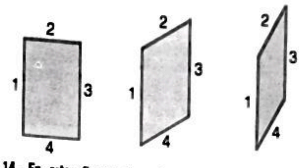
11. Los ángulos se miden con trasportador.



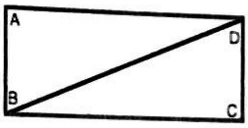
12. Arriba se muestran las tres clases de triángulos.



13. En a se ven tres segmentos de recta (1, 2, y 3). En b estos segmentos se han unido para formar un triángulo

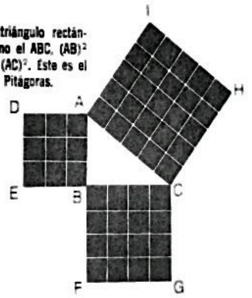


14. En estas figuras de cuatro lados, los marcados 1 son todos iguales, y lo mismo los marcados 2, 3 y 4.

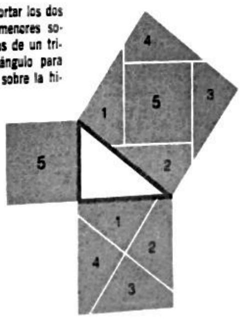


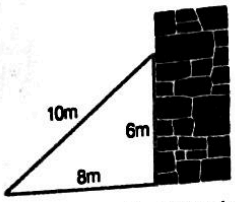
15. Para reforzar la construcción de cuatro lados se inserta la pieza diagonal BD, trasformando a la primera en dos triángulos rígidos.

16. En un triángulo rectángulo, tal como el ABC, $(AB)^2 + (BC)^2 = (AC)^2$. Este es el teorema de Pitágoras.

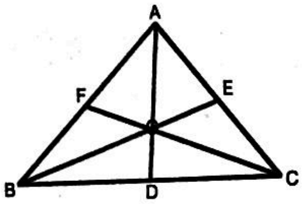


17. Cómo cortar los dos cuadrados menores sobre los lados de un triángulo rectángulo para formar uno sobre la hipotenusa.

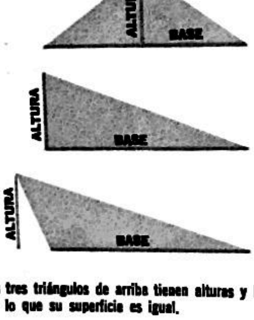




18. Si la pared es perpendicular al piso, el triángulo encajará cómodamente.

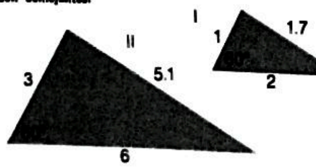


. F, E y D son los puntos medios de los lados AB, AC BC. Obsérvese que las tres líneas FC, BE y AD se cortan G. El texto señala otros aspectos interesantes de G.



20. Los tres triángulos de arriba tienen alturas y bases iguales, por lo que su superficie es igual.

21. Los triángulos con ángulos correspondientes iguales son semejantes.



caciones prácticas; el carpintero lo emplea para ver si una pared es perpendicular al suelo. Por ejemplo, si se juntan tablas de 6, 8 y 10 metros del modo indicado en la Figura 18, el ángulo entre las dos longitudes de 6 y 8 metros debe ser recto, pues $6^2 + 8^2 = 10^2$, o bien, $36 + 64 = 100$. Si la pared es verdaderamente perpendicular al suelo, el triángulo encajará exactamente.

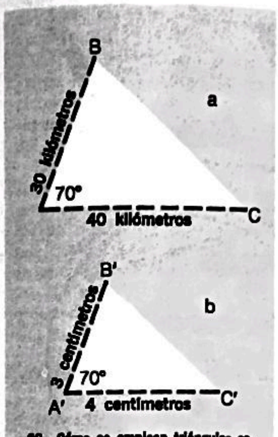
El teorema de Pitágoras es uno de los muchos que revelan una relación inesperada e importante. He aquí otro caso de tal teorema: en el triángulo *ABC*, de la Figura 19, *D*, *E* y *F* son los puntos medios de los lados *BC*, *AC* y *AB*. Los puntos *A*, *B* y *C* son los vértices del triángulo. Si unimos el *A* con el punto medio D del lado opuesto –*BC*–, la línea *AD* se denomina ***mediana***. Tracemos las dos medianas restantes del triángulo, *CF* y *BE*. Nos enteramos de lo que no habíamos sospechado: que las tres medianas pasan por el mismo punto *G*, interior al triángulo. También descubrimos que *G*, en cualquiera de las medianas, está a dos tercios de la distancia desde el vértice al lado opuesto. La geometría plana ofrece pruebas de estos asertos. He aquí otro hecho interesante respecto a *G:* si recortamos un triángulo de cartulina y trazamos las tres medianas –como en la Figura 19– podemos equilibrar el triángulo sobre el extremo romo de un lápiz, si colocamos dicho extremo directamente debajo de *G*. Este punto se denomina ***centro de gravedad***.

TRIÁNGULOS IGUALES Y SEMEJANTES

Cuando los triángulos tienen el mismo tamaño, o superficie, se denominan ***triángulos iguales:*** todos los triángulos con alturas y bases iguales lo son, si bien pueden tener muchas formas diferentes (Figura 20).

Algunos tienen la misma forma, pero tamaños diferentes; se los conoce como ***triángulos semejantes***. Los ángulos correspondientes de éstos son iguales, y sus lados correspondientes son siempre proporcionales. Así, los dos triángulos I y II que se muestran en la Figura 21 son semejantes, porque los ángulos correspondientes son iguales. Cada lado del triángulo II es el triple del correspondiente del triángulo I.

Se usan triángulos semejantes en el dibujo a escala; al hacer un mapa, por ejemplo, representamos una gran superficie de terreno en un pequeño trozo de papel. Supongamos que el lugar *A*, en la Figura 22*a*, esté a 30 kilómetros del lugar *B* y a 40 del *C*, y que el ángulo *CAB* es de 70°. Hemos de señalar estos lugares en un mapa en el que la escala ha de ser 1 centímetro = 10.000 metros. Primero, usando un trasportador, trazamos un ángulo de 70°. Treinta kilómetros son iguales a 30.000 metros, pues hay 1.000 metros en un kilómetro. Como un centímetro es igual a 10.000 metros, para representar 30 kilómetros –o sea 30.000 metros– precisamos tres centímetros; para 40 kilómetros, 4 centímetros. Medimos tres centímetros sobre un lado del ángulo de 70°, y cuatro sobre el otro. Al unir los extremos de estos segmentos tenemos un mapa (Figura 22b) que representa el área triangular formada por los tres lugares.

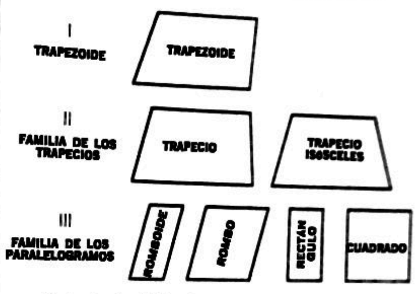


22. Cómo se emplean triángulos semejantes al elaborar mapas y para dibujar a escala.

Todos los mapas, modelos y fotografías son semejantes a los objetos originales que representan; por eso, los ángulos de estas representaciones son exactamente de la misma medida que en el original, y todas las líneas se modifican en la misma proporción. Sus superficies también tendrán una proporción definida: variarán como el cuadrado de los lados correspondientes. Si duplicamos cada uno de los lados de un triángulo, el área será el cuádruplo; si triplicamos cada lado, la superficie será 9 veces mayor. La Figura 23 demuestra la verdad de esto.

23. Si duplicamos la longitud de los lados del triángulo I, produciendo el II, la superficie de este último será cuatro veces mayor que la del I. La superficie del triángulo III, arriba, es nueve veces mayor que la del I.

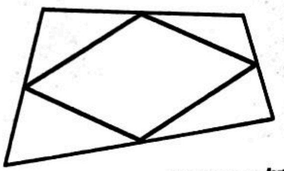
Si se proyecta película de un centímetro cuadrado sobre una pantalla, de modo que la imagen que aparece sea de 40 centímetros cuadrados, la proyección es 40^2, o sea 1.600 veces más grande que el original. La luz empleada para proyectar la película debe cubrir una superficie 1.600 veces mayor que la película, de donde se explica que su intensidad sobre la pantalla sólo sea 1/1.600 que en la película. Ésta es una ilustración notable de la manera en que puede aplicarse la geometría de las figuras semejantes al estudio de fenómenos fotográficos.



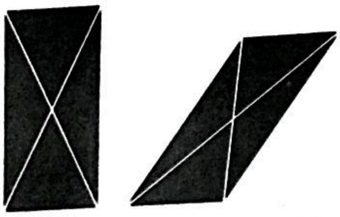
24. Los tres tipos básicos de cuadriláteros, o figuras de cuatro lados.

CUADRILÁTEROS

Una figura de cuatro lados se llama ***cuadrilátero***. Esta familia se muestra en la Figura 24. Al examinar los cuadriláteros vemos que, en realidad, hay tres tipos básicos: el ***trapecio***, con un par de lados paralelos; el ***trapezoide***, que no tiene lados paralelos, y el



25. En un cuadrilátero, la figura interna que se forma uniendo los puntos medios de los lados es un paralelogramo.

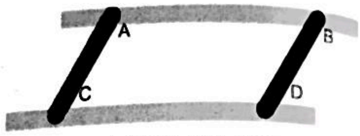


26. Los lados correspondientes de estos paralelogramos son iguales. Las diagonales de ambos se bisecan entre sí.

paralelogramo, con dos pares de lados paralelos. El *trapecio isósceles* es aquel cuyos dos lados no paralelos son iguales. Un *romboide* es un paralelogramo sin ángulos rectos. Un *rombo* es un romboide con los cuatro lados iguales. Un *rectángulo* es un paralelogramo cuyos ángulos son todos rectos. Un *cuadrado* es un rectángulo de lados iguales. En todas estas figuras, si unimos los puntos medios de los lados –como en la Figura 25– el cuadrilátero interno que se forma siempre será un paralelogramo.

En esta última figura, las diagonales se bisecan entre sí, no importa cómo la distorsionemos (Figura 26): éste es un buen ejemplo de *invariante:* propiedad de la figura que permanece cierta bajo todas las distorsiones. Otro invariante es: "Los lados opuestos de un paralelogramo son iguales."

El dibujante hace uso de este invariante en el instrumento llamado reglas paralelas (Figura 27): consiste en dos reglas unidas por dos varillas AC y BD, de modo y forma tales que $AC = BD$. Este dispositivo es flexible, por lo que AB puede hallarse a distancias variables de CD. Empero, no importa cuánto se aleje



27. El instrumento llamado reglas paralelas.

28. Arriba se muestran las partes principales de un círculo.

AB, siempre permanece paralela a CD. De este modo, mediante este adminículo podemos trazar una paralela a una recta dada, en cualquier punto accesible del plano.

El estudio de triángulos y cuadriláteros constituye una parte sustancial de la geometría. Todos los otros polígonos –figuras cerradas que tienen tres o más ángulos y, en consecuencia, lados– pueden dividirse en triángulos y cuadriláteros trazando diagonales desde los vértices del polígono; esto significa, pues, que las reglas básicas para determinar el perímetro y la superficie de triángulos y cuadriláteros pueden aplicarse al estudio de los otros polígonos.

EL CÍRCULO

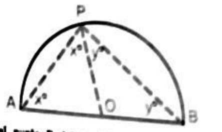
También es importante el estudio del círculo. En la Figura 28 mostramos sus partes principales; se observará que cuanto más cerca está una cuerda del centro del círculo, tanto más grande se vuelve. Realmente, el *diámetro* es una cuerda que pasa por el centro: consiste en dos radios unidos de modo de formar una recta. Para medir la longitud de un círculo –esto es, su *circunferencia*– hallamos cuántos diámetros caben en ella; el número es, aproximadamente, 3 1/7 o con más precisión, 3.1416; su valor exacto se indica con la letra griega π ("pi"). Luego, la circunferencia del círculo es πd, o sea, π veces el diámetro. Su superficie es πr^2, o sea π veces el cuadrado del radio.

La mitad de un círculo se denomina *semicírculo*. Un teorema simple y bastante sorprendente con respecto a un semicírculo es: "Si un punto cualquiera de un semicírculo se une con los extremos del diámetro, se forma en ese punto un ángulo de 90°." En la Figura 29, AB es el diámetro y P es un punto cualquiera del semicírculo. Es fácil probar que el ángulo APB tiene 90°, o sea, es recto. Primero, unamos P al centro del círculo (O); AO, PO y OB son todos radios del círculo, y por tanto, iguales. En el triángulo APO, como el lado $AO = PO$, los dos ángulos marcados $x°$ son iguales, pues si dos lados de un triángulo son iguales, los ángulos opuestos también deben serlo. Análogamente, para el triángulo POB: los lados OP y OB son iguales, por lo que los ángulos marcados $y°$ deben serlo también. Si dejamos de lado la línea OP, tenemos el triángulo APB, cuyos ángulos deben totalizar 180°, puesto que la suma de los ángulos de un triángulo es de 180°. Hay dos ángulos $x°$, y dos $y°$ en el triángulo APB; luego un ángulo $x°$ y un $y°$ deben dar la mitad de 180°, o sea 90°. Como el ángulo APB está compuesto por un $x°$ y un $y°$, debe ser igual a 90°, o sea, un ángulo recto.

Este teorema tiene varias aplicaciones: por ejemplo, un fabricante de moldes puede determinar si la matriz que se muestra en la Figura 30 es un verdadero semicírculo. Coloca un instrumento llamado escuadra de carpintero en la matriz; si tiene buen contacto en tres puntos –como se muestra– sabe que el molde le dará un verdadero semicírculo.

Un ángulo en el centro de un círculo tiene tantos grados como el arco que corta en el círculo. 0, como diría un matemático, "Un ángulo central se mide por su arco interior". En la Figura 31, el ángulo obtuso en 0 es igual a 110°, y el arco AB que corta también lo es.

Un ángulo cuyo vértice esté en la circunferencia de un círculo y que corta un arco, se denomina *ángulo inscrito* (ángulo ABC de la Figura 32). No importa dónde coloquemos el

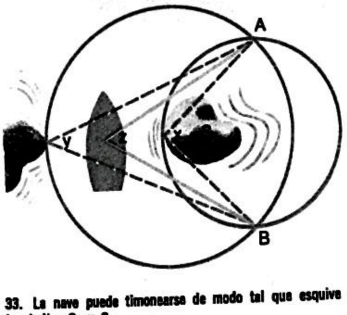


29. Si el punto P del semicírculo se une al diámetro en A y B, el ángulo APB será de 90°.

30. Cómo se determina, con una escuadra de carpintero, si una matriz da un semicírculo verdadero.

31. Tanto el ángulo en 0, como el arco que corta (AB) en el círculo de arriba, son iguales a 110°.

32. Los ángulos marcados 1, 2 y 3 cortan al mismo arco (AC) en este círculo: por tanto, son iguales.



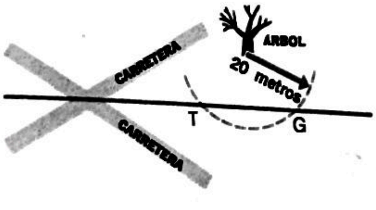
33. La nave puede timonearse de modo tal que esquive los bajíos S_1 y S_2.

vértice *B* respecto de este arco, el ángulo conservará la misma medida; en otras palabras, todos los ángulos inscritos que cortan el mismo arco son iguales.

Este invariante la emplean los navegantes para mantenerse alejados de los obstáculos. En la Figura 33, *A* y *B* son dos hitos, y S_1 y S_2, dos bajíos. Trazamos dos círculos, cada uno de los cuales tiene los puntos *A* y *B* en su circunferencia; el círculo menor no sólo pasará por *A* y *B*, sino que también contendrá el bajío llamado S_1; el otro, S_2, estará fuera del círculo grande.

Para mantenerse entre los bajíos, un barco debe permanecer fuera del círculo menor y dentro del mayor. En ángulo *x* siempre tiene la misma medida, doquiera esté en el círculo menor, pues siempre corta el mismo arco; por la misma razón, el ángulo *y* también tiene la misma medida. Mediante el sextante, el navegante trata de que, a medida que el buque navega entre S_1 y S_2, el ángulo *z* se mantenga menor que el *x* (lo que lo mantiene fuera del círculo menor), pero mayor que el *y* (lo que lo mantiene dentro del círculo mayor): así, el barco evita ambos bajíos.

34. Para hallar un tesoro a 20 metros de un árbol y a igual distancia de dos carreteras, úsese un locus.

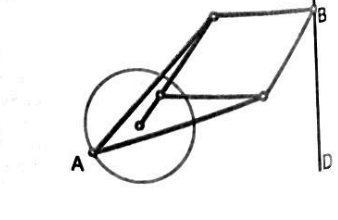


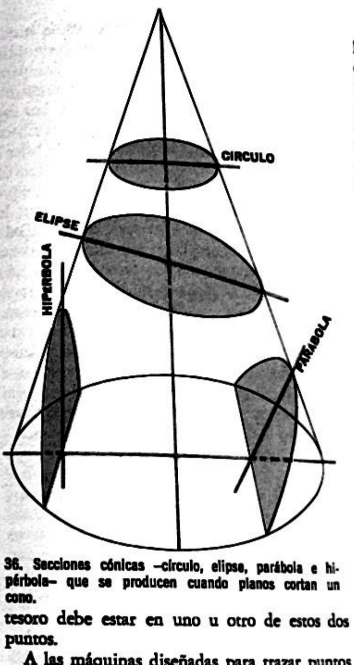
TRAYECTORIA DE UN PUNTO MÓVIL

A menudo, en geometría plana resulta necesario determinar la trayectoria que describe un punto en un plano cuando se mueve según una regla fija. Por ejemplo: si un punto debe permanecer siempre a 3 centímetros de otro fijo, viaja en círculo en torno de este último. El matemático da el nombre de ***locus*** a la trayectoria descrita por un punto. (*Locus* significa "posición" en latín). Por medio del estudio de las trayectorias de los puntos móviles determinamos cómo se mueven las partes de una máquina, y cómo parecen desplazarse los cuerpos celestes.

Podemos ilustrar el uso del locus mediante un problema muy simple de búsqueda del tesoro: se informa que, a 20 metros de un roble, y a igual distancia respecto de dos carreteras que se intersecan, hay un tesoro enterrado (Figura 34). Puesto que está a 20 metros del roble, se halla en algún lugar de la circunferencia de un círculo de 20 metros de radio con centro en el árbol. Ahora bien: también sabemos que el tesoro equidista de las dos carreteras que se intersecan. Podemos pensar en él como en un punto móvil que permanece a la misma distancia de cada carretera. En geometría se puede probar que un punto que se desplaza de forma tal que se mantiene igualmente distante de los lados de un ángulo traza una línea que biseca al ángulo. Por ello, el tesoro debe estar en la bisectriz del ángulo formado por las dos carreteras, la que corta al círculo –tal como se muestra en la Figura 34– en dos lugares, *T* y *G*. En consecuencia, el tesoro debe estar en uno u otro de estos dos puntos.

35. La celda de Peaucellier cambia el movimiento de circular a rectilíneo. Es un tipo de sistema articulado.





36. Secciones cónicas –círculo, elipse, parábola e hipérbola– que se producen cuando planos cortan un cono.

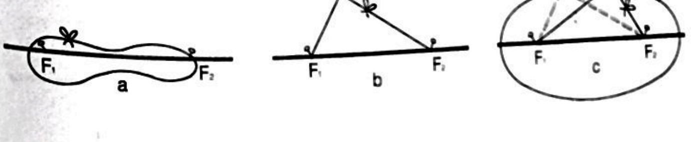
A las máquinas diseñadas para trazar puntos móviles se las llama *sistemas articulados*, porque consisten en barras articuladas: el más simple de estos sistemas es un par de compases. La trayectoria que traza es un círculo. Otro sistema articulado –llamado celda de Peaucellier– cambia el movimiento de circular a rectilíneo (Figura 35): a medida que *A*, en la figura, se mueve en torno del círculo, el punto *B* se desplaza hacia arriba y hacia abajo en la línea recta *CD*. Existen otros tipos de articulaciones que trasforman el movimiento circular en lineal. Fue necesario un estudio de sistemas articulados para ayudar a resolver el problema de conseguir funcionamiento suave en una locomotora, donde había que convertir el movimiento rectilíneo del vástago del émbolo en el movimiento circular de las ruedas.

SECCIONES CÓNICAS

El círculo es el tipo de curva más común, pero las hay de otras clases. Desde temprano se apercibieron los geómetras griegos de que, cuando se cortaba un cono con planos en ángulos diferentes, las intersecciones daban distintas clases de curvas: círculos, elipses, parábolas o hipérbolas (Figura 36). Puesto que las tres últimas clases de curvas se describieron por primera vez en conexión con conos, se las llamó *secciones cónicas*. Apolonio de Perga, que vivió en el siglo III a.C., escribió un tratado sobre las propiedades de estas curvas. En épocas más recientes se descubrió que también era posible definirlas como trayectorias descritas en el plano por puntos que se mueven según ciertas reglas. Tales definiciones son particularmente significativas cuando damos uso práctico a las curvas.

La elipse. Una elipse es la trayectoria trazada por un punto que se mueve de modo tal que la suma de sus distancias a dos puntos fijos siempre es la misma; los dos puntos fijos se denominan *focos*. Resulta fácil dibujar una elipse usando el método ilustrado en la Figura 37: primero, clavamos tachuelas en dos puntos fijos F_1 y F_2. Luego, tomamos un trozo de cuerda más largo que la distancia entre F_1 y F_2. Unimos un extremo de la cuerda a la tachuela F_1, y el otro a la F_2 (Figura 37*a*). Tensamos la cuerda e insertamos un lápiz tal como se muestra en la Figura 37*b*. A medida que mo-

37. Es fácil dibujar una elipse si se emplean dos tachuelas y una cuerda, como se muestra en los dibujos de abajo.



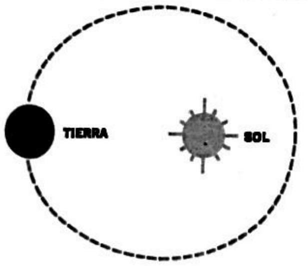
vemos el lápiz, su punta trazará una elipse (Figura 37*c*). La suma de las distancias desde la punta del lápiz móvil hasta los puntos fijos permanecerá constante: en la Figura 37*c*, $P_1F_1 + P_1F_2 = P_2F_1 + P_2F_2$.

Si una mesa de billar tuviese forma elíptica, cualquier bola arrojada desde un foco rebotaría hacia el otro. En un cuarto elíptico, cualquier sonido que se emita desde un foco será reflejado por las paredes al otro: éste es el principio de la "galería murmurante", un ejemplo de la cual esá dado por el tabernáculo mormón de forma elíptica, en Salt Lake City, Utah. Los focos del tabernáculo están marcados claramente; una persona que esté parada en un foco puede oír distintamente el murmullo de una persona que esté en el otro; los que están cerca, nada oyen.

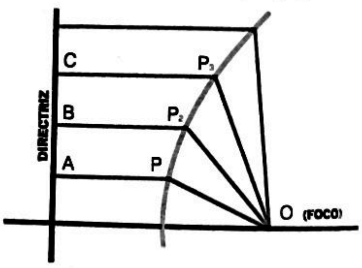
La elipse encontró muchas aplicaciones prácticas: las máquinas perforadoras eléctricas usan engranajes elípticos; en los extremos de la elipse, los engranajes se mueven más de prisa, dando un retorno más rápido; en las partes menos curvadas los engranajes se mueven más lentamente, ejerciendo mayor fuerza. Los tanques de almacenamiento y trasporte se hacen con sección trasversal elíptica, a fin de bajar el centro de gravedad y disminuir el peligro de vuelco.

La elipse también sirve para explicar los movimientos de varios cuerpos celestes. Todos los planetas se desplazan en órbitas elípticas, con el Sol en uno de los focos (Figura 38); Un planeta se desplaza en su órbita, de tal modo que el radio que va del Sol al planeta recorre superficies iguales en tiempos iguales; conociendo la órbita elíptica de cualquier planeta, los astrónomos pueden predecir su posición en su órbita en cualquier momento.

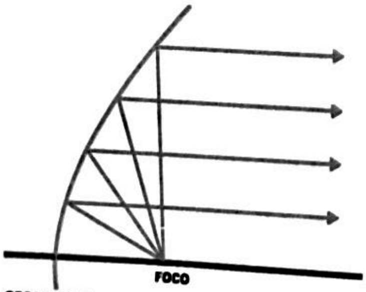
La parábola. La parábola es la trayectoria de un punto que se desplaza de forma tal que su distancia a una recta fija –llamada *directriz*– es siempre igual a su distancia a un punto fijo llamado foco. Así, en la Figura 39, a medida que un punto se mueve a lo largo de la parábola ocupando sucesivamente las posiciones P_1, P_2 y P_3 se tiene: $AP_1 = P_1O$; $BP_2 = P_2O$, $CP_3 = P_3O$. El reflector buscahuellas tiene una superficie parabólica con la fuente luminosa en el foco: todos los rayos de luz que emanan de éste son reflejados por la parábola en rayos paralelos (Figura 40). Los detectores sónicos tienen superficies parabólicas; las ondas sonoras se reflejan al chocar con la superficie y son concentradas en el foco. El espejo de un telescopio reflector tiene forma de parábola; los rayos paralelos de luz provenientes de un cuerpo celeste lejano inciden sobre la superficie parabólica y, reflejados por ella, se encuentran en el foco, dentro del tubo del telescopio.



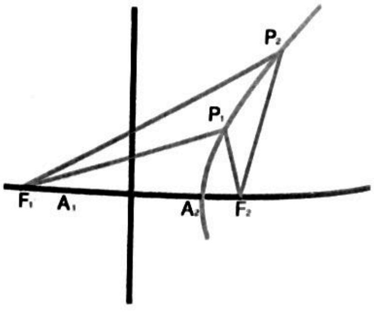
38. A medida que la Tierra viaja en torno del Sol recorre una órbita elíptica, con éste en uno de los focos.



39. En un punto móvil sobre una parábola, su distancia a la directriz y al foco (fijos) es igual.



40. Los haces luminosos que emanan del foco del buscahuellas, mostrado abajo, se reflejan desde la superficie parabólica en una serie de rayos paralelos.



41. P_1 y P_2 son dos posiciones de un punto que se desplaza sobre una hipérbola. La distancia a F_1 menos la distancia a F_2 siempre es igual a la distancia A_1A_2.

La hipérbola. La hipérbola es la trayectoria de un punto que se desplaza de tal forma que la distancia a un punto fijo menos la distancia a otro, también fijo, sea siempre la misma. El diagrama de la Figura 41 muestra una hipérbola en la que F_1 y F_2 –llamados focos– son los puntos fijos. P_1 es un punto de la hipérbola. P_1F_1 menos P_1F_2 es igual a A_1A_2. P_2 es otro punto de la hipérbola. P_2F_1 menos P_2F_2 también es igual a A_1A_2.

La hipérbola se aplica al LORAN, o navegación a larga distancia mediante el uso de radar. Podemos dar una explicación general del loran, refiriéndonos nuevamente a la Figura 41: hay dos estaciones de radar, una llamada estación maestra (F_2 en el diagrama), y la otra, esclava (F_1). Están en tierra, a unos 300 kilómetros de distancia entre sí. Cada estación emite pulsaciones eléctricas, las que tardan más tiempo para viajar de F_1 a P_1, que de F_2 a P_1, y desde F_1 a P_2, que desde F_2 a P_2. La diferencia en tiempo es una constante fija para todos los puntos de la hipérbola; si la diferencia de tiempo es mayor o menor, tenemos una hipérbola diferente.

Un avión, o un barco, tienen un instrumento para recepción de radar que recoge esta diferencia de tiempo entre las pulsaciones provenientes de las dos estaciones F_1 y F_2. El navegante consulta un mapa en que se han trazado las diversas hipérbolas correspondientes a las varias diferencias de tiempo, y localiza su propia nave en la hipérbola; entonces escoge la que pasa por el "puerto de origen" y anota la diferencia de tiempo de las pulsaciones. Se modifica el curso, hasta que el receptor de radar indica que la embarcación está en la hipérbola del "puerto de origen". Manteniendo constante la diferencia de pulsación, la nave va a su destino a lo largo de la hipérbola.

A edad muy temprana se enseña [illegible] las formas geométricas básicas que [illegible] que los rodea

GEOMETRÍA DEL ESPACIO

El mundo bidimensional de la geometría plana no es suficiente para explicar el mundo en que vivimos, pues éste posee tres dimensiones. En él hay muchos planos, que son ilimitados y se extienden en todas las direcciones imaginables. También hay muchos tipos de superficies curvas. Debemos considerar no sólo norte, sur, este y oeste, sino también arriba y abajo. Para explicar este mundo tridimensional se ha desarrollado el ramo de las matemáticas denominado geometría del espacio. Usamos este tipo de geometría en la construcción de maquinarias, rascacielos, aeroplanos, vapores, puentes y automóviles, y también para explicar los fenómenos del Universo.

En la geometría del espacio hay muchas más relaciones posibles entre elementos geométricos que en la geometría plana. En un solo plano, dos líneas son paralelas o se intersecan. También, en la geometría del espacio, dos líneas pueden ser paralelas o intersecarse, pero también pueden ser líneas ***alabeadas***. Las líneas alabeadas no pertenecen al mismo plano, nunca son paralelas y nunca se intersecan. En el plano, puede trazarse sólo una línea perpendicular a otra por un punto dado. En geometría del espacio puede trazarse cualquier número de dichas líneas perpendiculares. Por ejemplo, los rayos de una rueda son rectas, cada una de las cuales es perpendicular al eje en el mismo punto. En un plano, todos los puntos que equidistan de otro fijo están sobre un círculo. En el espacio tridimensional, sin embargo, están sobre una esfera que contiene un número infinito de círculos, que pasan por el centro.

ÁNGULOS EN GEOMETRÍA DEL ESPACIO

El ángulo más simple en geometría del espacio se denomina ***ángulo diedro*** ("dos caras"): está formado por dos planos que se intersecan. La magnitud de este ángulo se mide por medio del ángulo plano, que está formado por dos líneas, una en cada cara, que cortan la arista, o intersección de los dos planos, en ángulos rectos (Figura 1). Cuando un avión ladea las alas, el ángulo de ladeo es un ángulo diedro entre la horizontal y la posición inclinada de las alas (Figura 2). El ángulo diedro se mide con un instrumento en el avión y la medida de este ángulo determina en parte la velocidad con que el aeroplano cambiará su dirección de ruta.

Cuando tres planos se encuentran en un punto forman un ***ángulo triedro*** (Figura 3). Cada uno de los ángulos que constituyen un ángulo triedro se denomina ***ángulo de cara***. En la Figura 3, *ADC*, *CDB* y *ADB* son todos ángulos de cara. Si más de tres planos se encuentran en un punto, el ángulo se llama ***ángulo poliedro*** (muchas caras). La suma de los ángulos de cara de un ángulo poliedro debe ser menor de 360°. A medida que la suma de los ángulos se aproxima a 360°, el ángulo se hace menos agudo, hasta que, en 360°, se convierte en un plano (Figura 4). Los cristales de los minerales muestran muchas clases de ángulos poliedros. Un análisis de estos ángulos hace posible identificar los distintos minerales.

CINCO CUERPOS BÁSICOS

La mayor parte del estudio de la geometría del espacio se basa en cinco cuerpos básicos: el prisma, el cilindro, la pirámide, el cono y la esfera.

El prisma. En un prisma, todos los lados, o superficies, de las caras son paralelogramos; las bases son polígonos paralelos e iguales. En la Figura 5 se muestran varios tipos de prismas. El más común es el cubo, cuyas caras son cuadrados (Figura 5*b*). En un prisma triangular (Figura 5*a*), las bases son triángulos. Se usan prismas triangulares de cristal al estudiar la refracción, o desviación, de la luz.

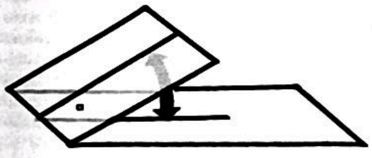
Las propiedades más importantes de un prisma son su área, o medida de la superficie, y su volumen, o medida del espacio que ocupa. El área lateral, o área de las caras, es igual al perímetro –o longitud del contorno– de la base multiplicado por la altura. Por supuesto, para hallar el área total del prisma se suman las áreas de las dos bases al área lateral. El interior de una habitación es generalmente un prisma. Si queremos decorar las paredes, por ejemplo, podemos hallar su área total tomando primero el perímetro del piso y multiplicándolo luego por la altura de la habitación. Si queremos pintar las paredes y el techo, hallamos el área lateral de la habitación y luego la del piso (que, en las habitaciones corrientes, es igual al área del techo). Luego sumamos las dos. Mediante dichos cálculos, un pintor o decorador estima el costo de los materiales para el trabajo que habrá de realizar.

El volumen de un prisma se halla multiplicando el área de la base por la altura. Este cálculo es muy importante en la construcción de una casa. La mayoría de los constructores estima el costo de la construcción a tanto por unidad cúbica. Para estimar cuánto costará construir una casa, primero hallaremos el volumen total de los prismas que constituirán la casa. Si el volumen buscado es de 800 metros cúbicos, y el constructor da una estimación de $ 30 por metro cúbico, el costo será $800 \times \$ 30$, ó $ 24.000.

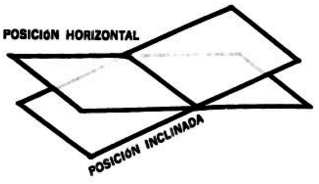
El cilindro. Si rotamos completamente un rectángulo sobre uno de sus lados, como en la Figura 6, definirá los límites del cuerpo denominado cilindro. Un envase de lata común es un buen ejemplo de cilindro: un cilindro circular recto, en el que las bases son círculos y los lados son perpendiculares a las bases. Sin embargo, no es el único tipo de cilindro, pues sus bases pueden ser elípticas y los lados no siempre perpendiculares a las bases.

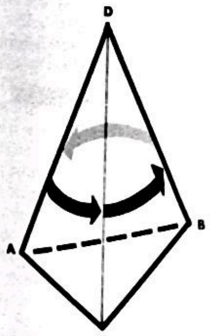
El área lateral de un cilindro circular recto es $2\pi ra$, en donde π es aproximadamente 3,1416, r es el radio de la base, y a, la altura. Supongamos que una compañía de envases decide fabricar una cantidad de envases de litro y quiere saber cuánto metal necesitará.

1. Los planos que se intersecan abajo forman un ángulo diedro. *a* es el ángulo plano (el ángulo entre los planos).



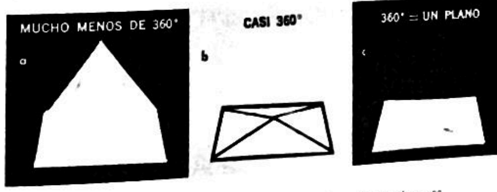
2. El ángulo diedro entre la horizontal y una posición inclinada del ala de un avión es el ángulo de inclinación.



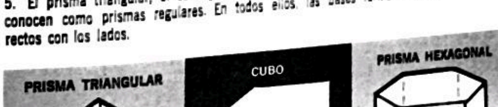


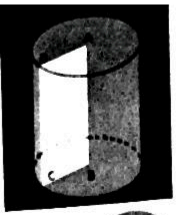
3. Cuando tres planos se encuentran en un punto, forman un ángulo triedro, como el que se muestra arriba.

4. Una serie de ángulos poliedros. A medida que la suma de los ángulos de las caras se hace mayor, el poliedro se vuelve menos y menos puntiagudo. A los 360° es un plano.

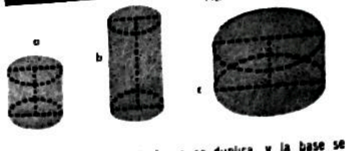


5. El prisma triangular, el cubo y el prisma hexagonal, mostrados abajo, se conocen como prismas regulares. En todos ellos, las bases forman ángulos rectos con los lados.





6 Si el **rectángulo** *ABCD* se rota alrededor del lado *AB*, marcará los límites de un cilindro, que tendrá a *AB* como eje



Si la altura de un cilindro *a* se duplica, y la base se mantiene igual, como en *b*, el volumen se duplica. Si la altura se mantiene igual, y el diámetro se duplica, como en c, el volumen del cilindro se cuadruplica.

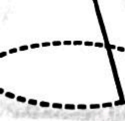
Esto se halla sumando las áreas de los dos círculos que constituyen las bases al área lateral.

El volumen de un cilindro se obtiene multiplicando el área de la base por la altura. La fórmula es $\pi r2a$, en la que *r* es el radio de la base y *a* es la altura. Es importante hallar el volumen de un cilindro al calcular el contenido de envases de lata, tanques de almacenamiento de gas, depósitos y demás, y también al determinar la velocidad de flujo y presión en tubos que conducen líquidos.

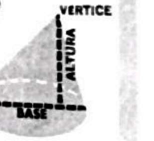
Si la altura del cilindro se duplica, y el diámetro se mantiene igual, su volumen también se duplicará. Se puede comprobar esto si se sitúa una lata al lado de otra exactamente igual. Sin embargo, si la altura de un cilindro se mantiene igual y el diámetro de la base se duplica, su volumen aumentará el cuádruplo (Figura 7).

Para ver si el tamaño económico de un producto vendido en latas tiene realmente un precio ventajoso, sería muy útil calcular el volumen del tamaño corriente y del tamaño económico y luego comparar los dos. Supongamos que la lata de tamaño corriente es de 12 centímetros de altura y tiene una base con un diámetro de 16 centímetros. Supongamos que la lata de tamaño económico es también de 12 centímetros de altura y tiene una base con un diámetro de 20 centímetros. El tamaño corriente vale $ 1,00 y el tamaño económico $ 1,25. El problema es ¿ahorraremos dinero si compramos el tamaño económico?

Sabemos que el volumen de un cilindro es $\pi r2a$. Para hallar el volumen de la lata corriente sustituimos *r* por 8 (el radio es la mitad del diámetro) y *a* por 12. Usando 3,14 como el valor de π, tendríamos $3{,}14 \times 64 \times 12 = 2.412$ centímetros cúbicos. Éste es el volumen de la lata común. El volumen de la lata de tamaño económico sería $3{,}14 \times 100 \times 12 = 3.768$ centímetros cúbicos. La lata de tamaño económico tendría alrededor de 1 vez y ½ el tamaño de una lata corriente ($3.768 \div 2.412$) y costaría una vez y ¼ lo que la otra. Representaría entonces un verdadero ahorro.



8 Pirámide cuadrangular. Tiene base cuadrada

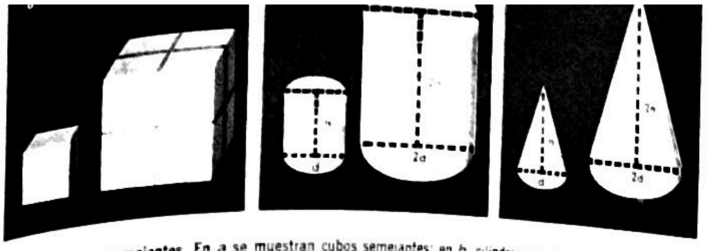


9 Tronco de pirámide cuadrangular

...se mantiene fijo mientras... sigue un camino circular. Se forma un cono...

11. Dos tipos de conos: a, cono circular recto; b, ...

12. Arriba, tronco de cono



13. Cuerpos semejantes. En *a* se muestran cubos semejantes; en *b*, cilindros semejantes; en c, conos semejantes.

La pirámide y el tronco de pirámide. En el cuerpo llamado pirámide, las caras son triángulos cuyos vértices se encuentran en un punto común y cuyas bases forman un plano (Figura 8). Si la base es un cuadrado, tenemos una pirámide cuadrangular. Las grandes pirámides de Egipto son cuadrangulares, así como los obeliscos denominados "agujas de Cleopatra".

Cuando el vértice de la pirámide se corta con un plano paralelo a la base, la parte inferior se denomina *tronco* (Figura 9). Las tiendas de campaña de un escuadrón del ejército y las tolvas de carbón, entre otras cosas, tienen la forma de tronco de pirámide. Para calcular el volumen de un tronco se requiere una fórmula bastante compleja. Aun hay evidencia de que los antiguos egipcios tenían una fórmula exacta para realizar dicho cálculo y la usaban a fin de determinar la cantidad de granito requerido para construir diferentes secciones de sus pirámides.

El cono y el tronco de cono. Un cono se forma sosteniendo un extremo de una línea fija y rotándola, según una trayectoria circular (Figura 10). Si el extremo fijo de la línea se sostiene directamente por encima del centro del círculo, el cono se denomina cono circular recto (Figura 11*a*). Si el vértice de un cono (punto más alto) no está directamente por encima del centro del círculo, el cono se llama oblicuo (Figura 11*b*). La altura de un cono es la distancia perpendicular del vértice a la base (Figura 11). Su volumen es igual a un tercio del producto del área de la base por la altura.

Si dividimos un cono en dos partes, atravesándolo con un plano paralelo a la base, la parte inferior se denomina tronco (Figura 12).

Muchas partes de máquinas tienen forma de conos o troncos de conos.

CARACTERISTICAS DE CUERPOS SEMEJANTES

Los cuerpos de la misma forma, pero de distintos tamaños, se llaman *semejantes*. Los ángulos poliedros correspondientes de los cuerpos semejantes son iguales, y las aristas correspondientes son proporcionales.

Las áreas de cuerpos semejantes tienen la misma relación que los cuadrados de las partes lineales correspondientes (ver Figura 13). En cada uno de los conjuntos de figuras semejantes en la Figura 13, la superficie de la figura mayor es cuatro veces la superficie de la menor, porque las partes lineales son dos veces más grandes. Si las partes se incrementaran tres veces, el área sería nueve veces mayor.

El volumen de cuerpos semejantes tiene la misma relación que los cubos de las partes lineales. En los conjuntos de cuerpos semejantes de la Figura 13, la figura mayor tiene ocho veces el contenido de la menor. En el caso de los cubos de la Figura 13*a* se pueden contar ocho cubos pequeños en el cubo mayor.

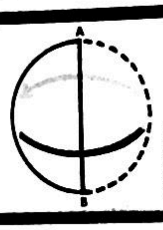
La razón de las áreas y volúmenes de cuerpos semejantes tiene muchas aplicaciones prácticas. Todas las esferas son semejantes. Si se venden naranjas de 8 centímetros de diámetro a 30 centavos la docena, mientras que las del mismo tipo y de 10 centímetros de diámetro se venden a 50 centavos, ¿cuál sería mejor comprar? El volumen de una naranja grande es $\left(\frac{10}{8}\right)^3$, ó $\frac{1.000}{512}$, o alrededor de dos veces el volumen de una naranja pequeña. Evidentemente, en este caso, las grandes serían una compra mucho mejor, porque cuestan menos del doble que las naranjas más pequeñas, pero tienen el doble de volumen.

La esfera. Si un semicírculo se rota alrededor de su diámetro (Figura 14), el cuerpo definido de ese modo se denomina esfera.

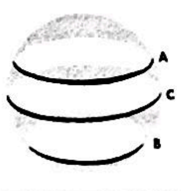
Cuando la esfera se corta con un plano, la intersección es un círculo. La Figura 15 muestra varios círculos formados de esa manera. Si el plano pasa por el centro de la esfera, el círculo de intersección tiene el mismo radio que la esfera. Ese círculo se denomina círculo máximo. Los demás son círculos menores.

La Tierra puede considerarse como una esfera en la que los polos Norte y Sur son los extremos de un diámetro denominado eje (Figura 16). Los círculos que pasan a través de ambos polos, el Norte y el Sur, son círculos máximos: se conocen como círculos de longitud. Todos los planos que cortan a la Tierra (excepto uno) en ángulo recto con el eje forman círculos menores denominados círculos de latitud. Hay sólo un plano que pasa por el centro de la Tierra y en ángulo recto con el eje. Forma un círculo máximo denominado ecuador.

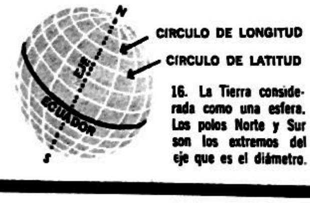
Entre dos puntos cualesquiera de la Tierra (sin incluir los polos) sólo puede dibujarse



14. A medida que un semicírculo rota alrededor de *AB*, que es el diámetro, se forma una superficie esférica.

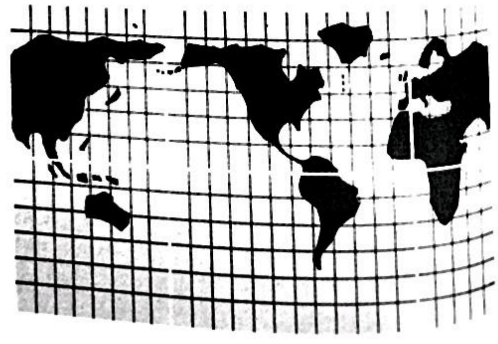


15. El plano C, que pasa por el centro de la esfera, es un círculo máximo. A y B son círculos menores.

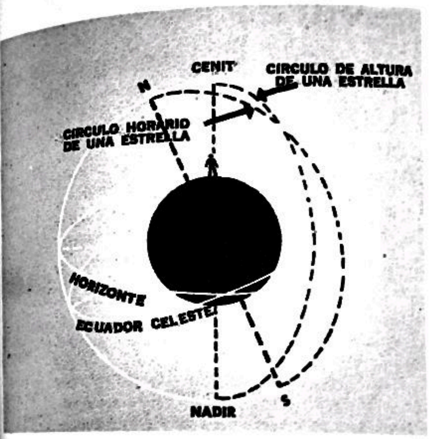


16. La Tierra considerada como una esfera. Los polos Norte y Sur son los extremos del eje que es el diámetro.

17. La menor distancia entre *A* y *B* en la esfera es el arco *AB*, que forma parte de un círculo máximo.



18. La proyección de Mercator. En dicho mapa, cualquier lugar de la Tierra puede localizarse determinando la latitud y la longitud del lugar.



19. Un diagrama simplificado que muestra la esfera celeste. La Tierra es el centro de esta esfera.

un círculo máximo (Figura 17). Todos los otros círculos que pasan por los dos puntos serán círculos menores. El arco más corto entre dos puntos es el arco de círculo máximo (*AB*, en la Figura 17). Los pilotos de aviones, dentro de lo posible, siguen el curso determinado por un arco de círculo máximo entre sus puntos de partida y destino.

Si consideramos la Tierra como una pelota de goma y la cortamos a lo largo de un semicírculo longitudinal, podemos estirar la pelota para formar una hoja rectangular plana. Los círculos de longitud se convertirán entonces en rectas verticales paralelas y los círculos de latitud, en rectas horizontales iguales y paralelas (Figura 18). Esta hoja representa ahora un mapa rectangular del mundo. Un mapa de este tipo se denomina una proyección de Mercator, por el geógrafo flamenco del siglo XVI Gerardus (o Gerhardus) Mercator, que la desarrolló. Al círculo máximo de longitud que pasa por Greenwich, Inglaterra, se le da el valor 0° de longitud; al ecuador, el 0° de latitud. En dicho mapa cualquier punto de la Tierra puede ubicarse determinando la longitud y la latitud. También, conforme dicho mapa, cuánto más nos alejamos del ecuador, tanto más estirada hallamos el área original, de modo que, en el mapa, las áreas terrestres cerca de los polos parecen mucho más grandes de lo que realmente son en la Tierra. Los que usan dicho mapa deben tener en cuenta tales distorsiones.

El área de una esfera es cuatro veces el área de un círculo máximo; la fórmula es $4\pi r^2$. El radio de la Tierra es aproximadamente 6.380 kilómetros; por tanto, la superficie total de la Tierra es $4 \times \pi \times 6.380^2$ kilómetros cuadrados, o alrededor de 511.000.000 de kilómetros cuadrados.

El volumen de una esfera puede expresarse por la fórmula $4/3\,\pi r^3$. Como el radio de la Tierra es 6.340 kilómetros, el volumen de nuestro planeta es $4/3 \times \pi \times 6.380^3$, lo cual será alrededor de 1.020.000.000.000 de kilómetros cúbicos.

USO EN ASTRONOMIA

La geometría del espacio ha permitido a los astrónomos dar una interpretación útil de los cielos y calcular la distancia y posición de los cuerpos celestes. El Universo se concibe como una inmensa esfera con un radio infinitamente grande, que parece rotar alrededor de la Tierra. En la Figura 19 hacemos una representación muy simplificada de dicha esfera como se la ve desde el punto de mira de un hombre ubicado a una latitud de 50°. Este hombre está en una esfera mucho más pequeña, que por supuesto es la Tierra. Directamente encima de su cabeza está el cenit; directamente debajo, el nadir. La línea donde el cielo parece encontrarse con la Tierra se denomina horizonte. Si el eje de la Tierra, que pasa por los polos Norte y Sur, se extiende, se encontrará con los límites exteriores de nuestra esfera celeste imaginaria en los polos celestes: norte y sur. La recta que une los dos polos celestes se denomina eje celeste. El plano del ecuador terrestre cortará los límites exteriores de la esfera celeste en un círculo máximo denominado ecuador celeste. Un círculo máximo que pasa a través de los polos y una estrella es el círculo horario de la estrella. El círculo de altura de la misma estrella es un círculo máximo que pasa por el cenit y la estrella.

Éstas son sólo algunas de las características de la esfera celeste. Ellas proporcionan una estructura de referencia que permite al astrónomo trazar los movimientos de los objetos del cielo, lo cual es una de las contribuciones destacadas de la geometría del espacio a la ciencia.

TRIGONOMETRÍA

Un importante ramo de la geometría es la trigonometría, o medida de triángulos. En trigonometría, cuando se conocen algunos datos de los triángulos, se puede determinar los elementos restantes, y entonces resolver gran variedad de problemas.

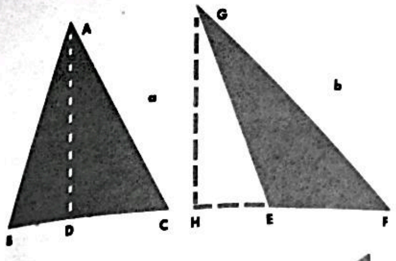
El fundador de la trigonometría fue el astrónomo griego Hiparco de Nicea, que vivió en el siglo II a.C. Hiparco intentó medir el tamaño del Sol y la Luna y sus distancias a la Tierra, y necesitó de un tipo de matemáticas que le permitiera medir objetos lejanos en el espacio, aplicando mediciones hechas sobre la Tierra. Esto lo llevó a la invención de la trigonometría.

Los boy scouts tienen ocasión de usar la trigonometría en su actividad campestre. Un problema común que se les plantea es calcular la altura de un árbol. El boy scout mide primero una distancia desde la base del árbol, digamos 20 metros, como se muestra en la Figura 1. Ésta será su "línea de base". En el punto *A*, mide el ángulo desde el suelo hasta la punta del árbol mediante un trasportador. Supongamos que este ángulo sea 35°. El boy scout conoce ahora dos datos del gran triángulo rectángulo formado cuando une los puntos *B* (la base del árbol), *A* (el extremo de la línea trazada desde el árbol) y *C* (la punta del árbol). Sabe que la longitud *AB* es de 20 metros y que el ángulo *BAC* es de 35°.

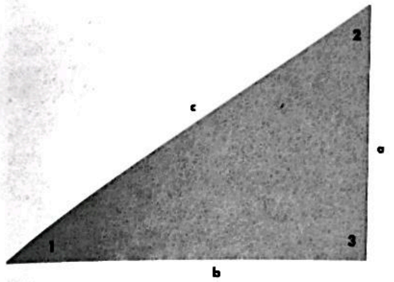
1. El problema es encontrar la altura del árbol cuando se conocen el ángulo en A (35°) y la distancia AB (20 metros). Mostramos cómo resolver el problema por medio de los triángulos semejantes ABC y A'B'C'.

Nuestro boy scout dibuja ahora sobre un papel un triángulo semejante al grande trazado en el terreno. Primero traza una línea *A'B'* de ½ metro de largo, y en *A'* dibuja un ángulo de 35°, –el ángulo *B'A'D'*– con un trasportador. Luego levanta por *B'* una perpendicular a la línea *A'B'*. Esta línea se intersecará con la *A'D'* en *C'*, y el ángulo *A'B'C'* será un ángulo recto. Los ángulos correspondientes de los triángulos grande y pequeño son iguales: ángulo *CAB* = ángulo *C'A'B'*; ángulo *ABC* = ángulo *A'B'C'*; ángulo *ACB* = ángulo *A'C'B'*. En consecuencia, tenemos dos triángulos semejantes, y los lados correspondientes serán proporcionales.

El boy scout mide ahora la línea *B'C'* y encuentra que es de 35 centímetros, o sea 0,35 metros. Dado que los lados de triángulos semejantes son proporcionales, *AB* es a *A'B'* como *BC* es a *B'C'*. Nosotros conocemos todas estas cantidades excepto *BC*, a la que podemos llamar *x*. Tenemos ahora la proporción 20 es a 0,5 como *x* es a 0,35, que la podemos escribir como $20 : 0{,}5 : : x : 0{,}35$. En cualquier proporción, el producto de los extremos (los dos términos exteriores) es igual al producto de los medios (los dos términos interiores); en consecuencia $x = 14$. La altura del árbol es, entonces, 14 metros. El boy scout resolvió este problema conociendo sólo un lado y un ángulo agudo de un triángulo rectángulo; usó los métodos básicos de la trigonometría, aunque un matemático, como veremos, no encararía el problema de semejante manera.



2. Cualquier triángulo puede convertirse en un triángulo rectángulo si se traza una perpendicular desde el vértice a la base.



3. Un triángulo rectángulo típico con la hipotenusa designada c, y catetos a y b.

FUNCIONES TRIGONOMÉTRICAS

La trigonometría está basada en el uso del triángulo rectángulo. Se puede aplicar a cualquier triángulo porque trazando una altura (una perpendicular desde el vértice a la base), podemos siempre convertirlo en triángulos rectángulos. En la Figura 2*a*, por ejemplo, la altura *AD* divide el triángulo *ABC* en los triángulos rectángulos *ADB* y *ADC*; en la Figura 2*b*, la altura *GH* convierte el triángulo *GEF* en los triángulos rectángulos *GHF* y *GHE*.

Ciertas razones básicas o relaciones entre los lados de un triángulo rectángulo son el corazón mismo del estudio de la trigonometría. Entre estas relaciones están el seno, el coseno, la tangente y la cotangente. Para entender qué significan estos términos, dibujemos un triángulo rectángulo típico, con ángulos 1, 2 y 3 y lados *a*, *b* y *c* (Figura 3). El ángulo 3 es recto; los otros dos ángulos son agudos (es decir, ángulos de menos de 90°). El lado *c*, que es el opuesto al ángulo recto, es la hipotenusa. Los otros dos lados, *a* y *b*, se llaman catetos. Podemos ahora definir el seno, coseno, tangente y cotangente como sigue:

El *seno* de cualquiera de los ángulos agudos es la razón del cateto opuesto a la hipotenusa. El seno del ángulo 1 es a/c; el del 2 es b/c.

El *coseno* de cualquiera de los ángulos agudos es la razón del cateto adyacente a la hipotenusa. El coseno del ángulo 1 es b/c; el coseno del ángulo 2 es a/c.

La *tangente* de cualquiera de los ángulos agudos es la razón del cateto opuesto al cateto adyacente. La tangente del ángulo 1 es a/b; la tangente del ángulo 2 es b/a.

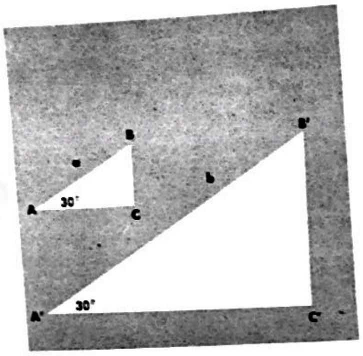
La *cotangente* de cualquiera de los ángulos agudos es la razón del cateto adyacente al cateto opuesto. La cotangente del ángulo 1 es b/a; la cotangente del ángulo 2 es a/b.

Se dice que el seno de un ángulo es una función trigonométrica de ese ángulo, porque su valor depende del tamaño del mismo. El coseno, la tangente y la cotangente también son funciones trigonométricas.

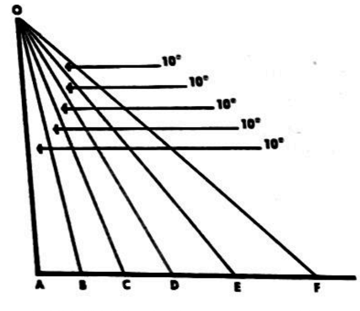
Una función trigonométrica dada, tal como el seno, es siempre la misma para un ángulo agudo dado de un triángulo rectángulo. En la Figura 4, por ejemplo, el ángulo *BAC* del triángulo rectángulo *ABC* es de 30°. Esto significa que el ángulo *ABC* debe ser de 60°, dado que el ángulo *ACB* es de 90° y la suma de los ángulos interiores de un triángulo es de 180°. El ángulo *B'A'C'* del triángulo rectángulo *A'B'C'* es de 30°, y el ángulo *A'B'C'* debe ser de 60°. En consecuencia, los ángulos correspondientes de ambos triángulos son iguales, y los lados correspondientes deben ser proporcionales. Por ello, $\frac{BC}{AB}$ (el seno del ángulo *BAC*, de 30°) y $\frac{B'C'}{A'B'}$ (el seno del ángulo *B'A'C'*, de 30°) deben ser iguales. En consecuencia, el seno de 30° es siempre el mismo, no importa cuán grande o pequeño sea el triángulo rectángulo en el cual se considere.

RESOLUCIÓN DE PROBLEMAS

Los matemáticos han calculado los valores de las funciones trigonométricas. Por ejemplo, el



4. El seno del ángulo BAC (30°) es igual al seno del ángulo B'A'C' en los dos triángulos semejantes.



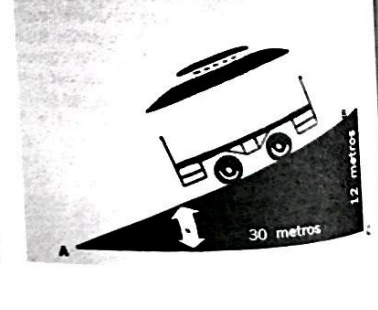
5. Este diagrama muestra cómo los matemáticos han determinado las tangentes de los ángulos.

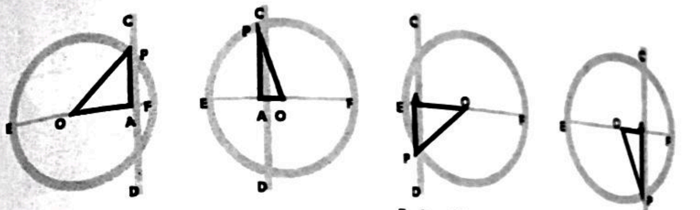
seno de 40°, con cinco decimales, es 0,64279; su coseno es 0,76604; su tangente es 0,83910; su cotangente es 1,1918. Para dar una idea de cómo se deducen esos números examinemos el procedimiento que se emplea para calcular las tangentes de diferentes ángulos agudos. Debemos recordar que la tangente de un ángulo es la razón del cateto opuesto al cateto adyacente.

En la Figura 5, el lado OA es exactamente 5 centímetros. Cada uno de los pequeños ángulos con vértice en O es exactamente de 10°. El ángulo BOA es, por tanto, de 10°; el ángulo COA, de 20°; el ángulo DOA, de 30°; el ángulo EOA, de 40°, y así sucesivamente. Medimos ahora AB y vemos que es 0,9 centímetros. Dado que la tangente del ángulo BOA es $\frac{AB}{OA}$ y que $OA = 5$, la tangente del ángulo es $\frac{0{,}9}{5}$, ó sea 0,18. Ésta es la tangente del ángulo de 10°, así tenga el lado OA centímetros, metros o kilómetros de largo. Midiendo por turno AC, AD, AE y así sucesivamente, podemos encontrar las tangentes de 20°, 30°, 40° y los demás ángulos.

Los valores de las funciones trigonométricas se encuentran en tablas especiales. Provistos con estas tablas es posible resolver con mucha facilidad una gran variedad de mediciones. Volvamos, por un momento, al problema del boy scout, como se presentó en la Figura 1a. Sabemos que AB es 20 metros y que el ángulo CAB es 35°. Debemos calcular BC (la altura del árbol). La tangente del ángulo CAB de 35° es $\frac{BC}{AB}$. Sabemos que AB es 20 metros; dado que BC es la cantidad desconocida, la llamamos x. En consecuencia, la tangente del ángulo CAB es $\frac{x}{20}$. Recorriendo la tabla de funciones trigonométricas, encontramos que la tangente de 35° es 0,7. Tenemos entonces la ecuación $\frac{x}{20} = 0{,}7$. Multiplicando ambos miembros de la ecuación por 20, tenemos $\frac{20x}{20} = 14$, luego $x = 14$.

6. El funicular se eleva 12 metros en una distancia horizontal de 30 metros. ¿Cuál es el ángulo en A?





7. A medida que el radio OP gira alrededor de la circunferencia en sentido contrario al de las agujas del reloj, forma una serie de triángulos con el diámetro de la circunferencia y con la varilla CD mantenida en posición vertical al extremo del radio.

En el problema precedente, la cantidad desconocida era un miembro de la razón de la tangente. Otros problemas trigonométricos pueden implicar la cotangente, el seno o el coseno. Además, en otros casos, la cantidad desconocida puede ser un ángulo, como en el problema siguiente.

Un funicular, en una colina de pendiente uniforme, sube 12 metros en una distancia horizontal de 30 metros. ¿Cuál es el ángulo de la pendiente, redondeado al grado más cercano? Esquematizamos el problema, como se ve en la Figura 6. Queremos encontrar el ángulo x. Sabemos que la tangente de x es $\frac{BC}{AC} = \frac{12}{30} = 0{,}4$. Consultando la tabla de funciones trigonométricas, vemos que 0,4 es la tangente del ángulo de 22° (redondeado al grado más cercano). Por tanto, el ángulo de la pendiente es 22°.

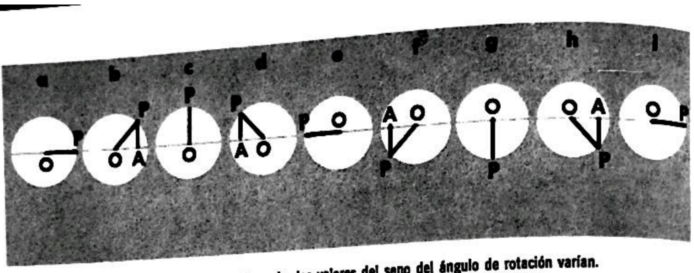
El agrimensor hace extenso uso de la trigonometría. Primero, trata de obtener una línea fija que no tenga obstrucciones para que pueda medirse con suficiente precisión. Ésta es la línea base que se usa en todos los cálculos. Las otras mediciones, en lo posible, son mediciones de ángulos. La trigonometría es también de vital importancia en la ingeniería, la navegación, la cartografía y la astronomía.

UTILIZADA PARA ESTUDIAR FENÓMENOS PERIÓDICOS

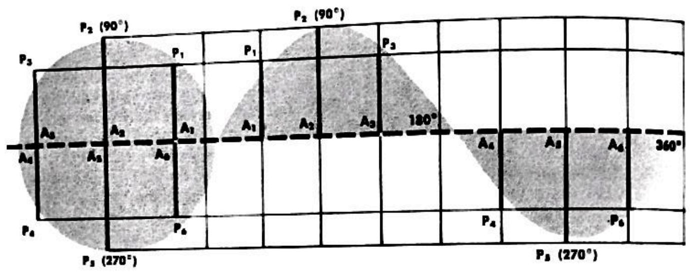
La trigonometría se usa además en otras formas. Una de ellas es el estudio de varios fenómenos periódicos. Cualquier fenómeno que se repite a intervalos de tiempo regulares se llama periódico. Las mareas, por ejemplo, son periódicas, dado que suben y bajan en una secuencia regular. El movimiento del plato de un péndulo también es periódico cuando oscila hacia uno y otro lado. Demostremos cómo describimos todos los fenómenos periódicos en función del seno de un ángulo.

Sabemos que el rayo de una rueda que gira recorre un ángulo de 360° al dar una vuelta completa. Repite el mismo recorrido en la segunda vuelta, en la tercera, y así sucesivamente. Evidentemente, tal rotación es periódica. El rayo de una rueda es en realidad el radio de una circunferencia. Podemos analizar el movimiento del radio alrededor del centro de la circunferencia si examinamos los diagramas en la Figura 7. Debemos suponer que una varilla, CD, se mantiene en posición vertical en el extremo del radio mientras éste se mueve alrededor del círculo. Se advertirá que se forma una serie de ángulos rectos al mantener la varilla su posición vertical. La línea AP, que une los puntos donde CD se encuentra con la circunferencia y con el diámetro EF, se alarga y luego se acorta. El ángula POA, que se llama ángulo de rotación, también cambia a medida que el radio gira alrededor del centro de la circunferencia. El seno de POA, el ángulo de rotación, es $\frac{AP}{OP}$. OP, la hipotenusa del triángulo rectángulo APO, es el radio de la circunferencia y, por supuesto, nunca cambia. Si le damos valor unitario (es decir, 1), AP representará el seno del ángulo de rotación, dado que el seno es igual a $\frac{AP}{OP}$.

Analicemos ahora los diferentes valores que toma el seno del ángulo de rotación a medida que el radio gira alrededor de la circunferencia. Los valores del seno situados por encima del diámetro se expresan como valores positivos. Los valores por debajo del diámetro son



8. A medida que el radio recorre la circunferencia, los valores del seno del ángulo de rotación varían.



9. Las variaciones del valor del seno se muestran aquí por medio de una gráfica. Los diferentes valores del seno PA se dan aquí como P_1A_1, P_2A_2, P_3A_3, etc. La línea curva a la derecha de la circunferencia se denomina curva del seno.

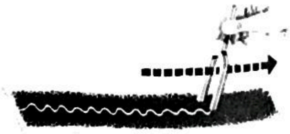
valores negativos. En la Figura 8*a*, el seno es cero, correspondiendo a un ángulo de rotación cero. El seno continúa creciendo a medida que el radio gira alrededor del centro (*b*) hasta que llega al valor 1 en *c*. Se hace luego más pequeño (*d*) hasta que llega a cero nuevamente en *e*. Después cruza por debajo del diámetro y toma valores negativos. Aumenta en tamaño (*f*) hasta que alcanza el valor — 1 en *g*. Luego, se vuelve más pequeño (*h*) hasta que, después que se ha producido una vuelta completa, se vuelve cero nuevamente (*i*).

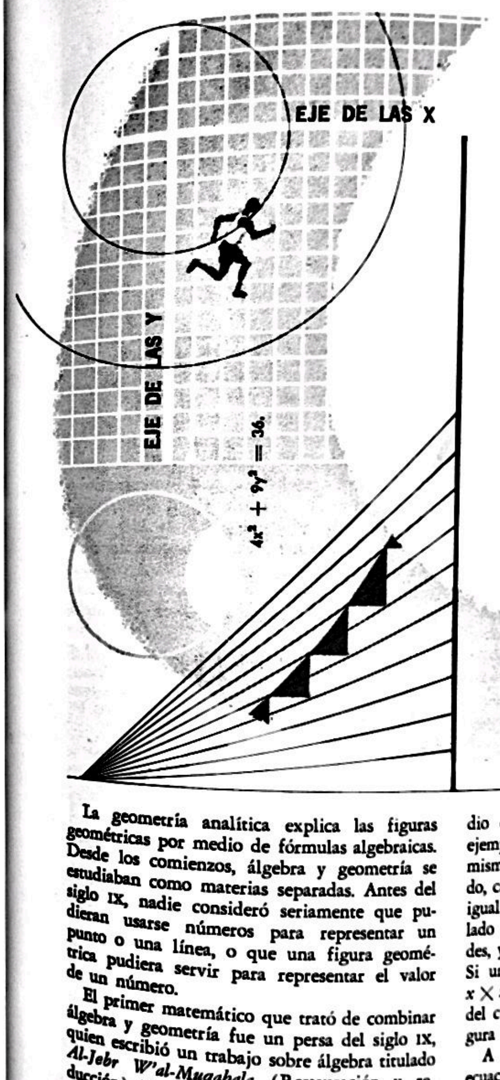
Podemos mostrar la variación del seno en la gráfica de la Figura 9. Damos los diferentes valores del seno *PA*, como P_1A_1, P_2A_2, P_3A_3, etcétera. La línea a la derecha del círculo se llama curva del seno. Se repite cada 360°.

La curva del seno puede aplicarse, entre otras cosas, a los fenómenos periódicos del sonido. Por ejemplo, cuando un diapasón se golpea, vibra, y la vibración causa ondas sonoras que se trasmiten al exterior. Si inmediatamente después de ser golpeado, un diapasón que da un tono de do intermedio se pasa muy rápidamente sobre una hoja de papel cubierta de hollín, la vibración del diapasón describe una serie de curvas de seno (Figura 10). Se producen doscientas sesenta y cuatro oscilaciones completas en un segundo; por tanto, el do intermedio corresponde a 264 vibraciones por segundo. Decimos que tiene una frecuencia de 264 hertz. Supongamos que golpeamos un diapasón que da un tono una octava más alto que el do intermedio y lo pasamos sobre un papel como antes. En este caso se producirán 528 sinusoides en un segundo. Esto significa que la frecuencia es el doble de la anterior.

Las curvas de seno de varias amplitudes y frecuencias se usan para explicar fenómenos eléctricos, luz (polarizada y simple) y fuerza, así como fenómenos sonoros. De este modo, la trigonometría ayuda a explicar y controlar nuestro medio físico.

10. Ondas sinusoidales producidas por un diapasón en vibración a medida que fue pasado rápidamente sobre una hoja de papel cubierta de hollín.



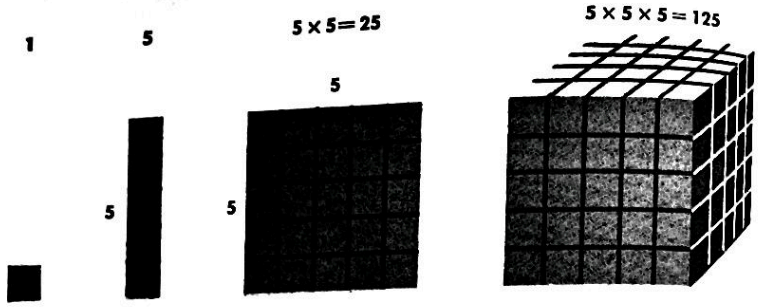


GEOMETRÍA ANALÍTICA

La geometría analítica explica las figuras geométricas por medio de fórmulas algebraicas. Desde los comienzos, álgebra y geometría se estudiaban como materias separadas. Antes del siglo IX, nadie consideró seriamente que pudieran usarse números para representar un punto o una línea, o que una figura geométrica pudiera servir para representar el valor de un número.

El primer matemático que trató de combinar álgebra y geometría fue un persa del siglo IX, quien escribió un trabajo sobre álgebra titulado *Al-Jebr W'al-Muqabala* (Restauración y reducción). En el mismo, este matemático mostró cómo resolver ecuaciones algebraicas por medio de cuadrados, rectángulos y cubos. Por ejemplo, para un número multiplicado por sí mismo –como en $x \times x$– recurría a un cuadrado, cada uno de cuyos lados tenía una longitud igual al valor (x). Si x fuese igual a 5, cada lado del cuadrado tendría un largo de 5 unidades, y éste estaría constituido por 25 unidades. Si un número fuese factor tres veces –como $x \times x \times x$–, él empleaba un cubo, cada lado del cual tenía una longitud igual a x. Ver Figura 1.

A causa de este método de resolución de ecuaciones, a un número multiplicado por sí mismo se lo denominó cuadrado de este número, en vez de potencia segunda. Por la



1. Para expresar un número multiplicado por sí mismo –digamos, 5 × 5– Mohamed ibn Musa usó un cuadrado, cada uno de cuyos lados era igual a ese número. Para expresar un número empleado tres veces como factor –como en 5 × 5 × 5– empleó un cubo.

misma razón, no leemos "x^3" como "x a la tercera potencia", sino como "x al cubo".

Mohamed ibn Musa no hizo esfuerzo alguno para resolver por el método geométrico una ecuación de cuarto grado, esto es, una ecuación con un término elevado a la cuarta potencia. Evidentemente, esto se debía a que ningún matemático de su tiempo tenía siquiera un atisbo de la cuarta dimensión. (Hoy en día, empero, la cuarta dimensión es un concepto vital y aceptado en las matemáticas y la física modernas.)

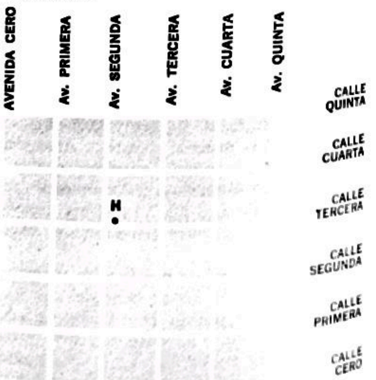
El esfuerzo pionero de Mohamed ibn Musa para quebrar las barreras entre el álgebra y la geometría fue infructuoso. Durante siglos, las dos materias siguieron estudiándose en compartimientos estancos, por así decirlo. Sólo al llegar el siglo XVII se derribó la partición entre aquellos compartimientos, y finalmente se desarrolló la geometría analítica.

SISTEMA DE COORDENADAS CARTESIANAS

La idea fundamental de la geometría analítica fue elaborada por el gran filósofo y matemático francés del siglo XVII René Descartes, quien afirmaba que la idea le sobrevino en un sueño. Presentó este nuevo enfoque de las matemáticas en su *Discurso del método*, en 1637. La idea básica de la geometría analítica es tan sencilla, que uno se pregunta cómo no se le ocurrió a ninguno de los matemáticos anteriores a Descartes.

Para explicarla, pensemos en una ciudad dispuesta en manzanas rectangulares, con las calles que van de este a oeste, y de norte a sur, tal como en la Figura 2; las calles de dirección este a oeste se denominarán Calle Cero, Calle Primera, Calle Segunda, Calle Tercera, y así sucesivamente. A las que van de norte a sur se las designará Avenida Cero, Avenida Primera, Avenida Segunda, Avenida Tercera, y así sucesivamente. Si las casas de cada manzana se numeran por centenas, será fácil localizar cualquier casa, siempre y cuando conozcamos el nombre de la calle y su número. Por ejemplo,

2. La casa que hemos indicado con *H* en este dibujo está en Calle Tercera 210; se halla 2,1 unidades al este, a partir de la Avenida Cero, contando cada cuadra como unidad. Está 3 unidades hacia el norte, empezando por la Calle Cero.

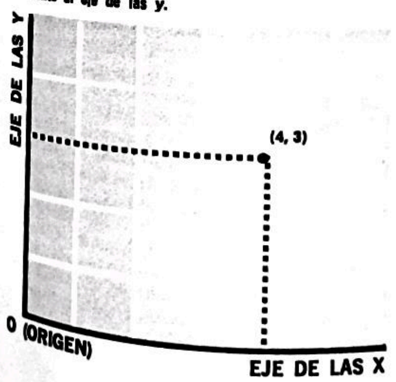


si hay una casa (*H*, en la Figura 2) en Calle Tercera 210, sabemos que está en la Tercera Calle, a un décimo de distancia desde la Avenida Segunda hasta la Tercera. Podríamos indicar la localización de la casa en forma más sencilla por medio de dos números: 2,1 y 3, escritos como (2,1, 3). 2,1 significaría 2,1 unidades hacia el este, contando cada cuadra como una unidad y empezando desde la Avenida Cero; el 3 significaría 3 unidades hacia el norte, empezando por la Calle Cero

La Calle Cero –que va de este a oeste en la ciudad que acabamos de describir– correspondería a lo que se denomina *eje de las x* en geometría analítica. Las distancias medidas sobre este eje se conocen como *distancias x*, o *abscisas*. La Avenida Cero, que corre de sur a norte, corresponde al *eje de las y*, y las distancias que se miden sobre él se llaman *distancias y*, u *ordenadas*. La intersección del eje x y del y se llama *origen*. Mostramos el eje de las x, el eje y las y y el origen (marcado con O) en la Figura 3

Un punto de un plano puede localizarse por medio de un par ordenado de números (x, y), con x siempre primero; x representa el número de unidades medido sobre el eje de las x, a partir del origen; y, el de unidades medidas sobre el eje de las y, a partir del origen; x y y se denominan *coordenadas cartesianas*. "Cartesiano" es el adjetivo correspondiente al nombre propio "Descartes".

3. Para localizar el punto (4,3), tomar 4 unidades hacia la derecha, a partir de O –el origen– sobre el eje de las x y luego 3 unidades hacia arriba, paralelamente al eje de las y.

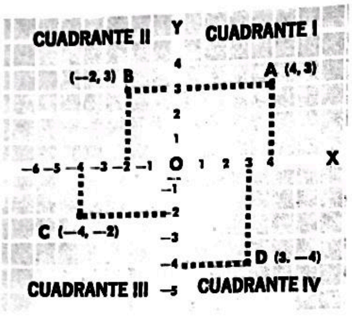


Un punto de un plano puede localizarse sustituyendo x y y por números particulares; así, uno puede hallar la posición del punto (4, 3) tomando 4 unidades hacia la derecha, desde el origen, a lo largo del eje de las x; y luego 3 hacia arriba, en forma paralela al de las y (Figura 3). Cualquier par de números, tales como (4, 3), localiza un punto. Cualquier punto puede representarse por un par de números. Más aún, valores diferentes de x y y representan puntos en todo un plano, los que, a su vez, describen varias figuras geométricas. Más adelante examinaremos algunas de estas figuras.

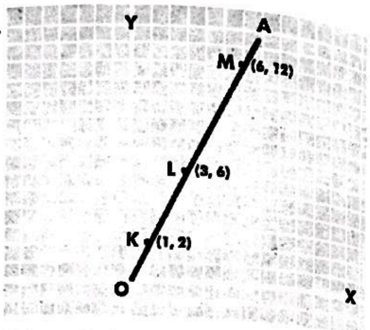
EXTENSIÓN DEL SISTEMA

En el sistema desarrollado por Descartes sólo se podía avanzar hacia la derecha del origen, sobre el eje de las x, y sólo hacia arriba –también desde el origen– sobre el de las y. Luego, los matemáticos extendieron el sistema continuando el eje de las x hacia la izquierda, más allá del origen (O), y haciendo lo propio con el eje de las y hacia abajo, más allá de O, como en la Figura 4. Éste es el sistema que se usa en la actualidad. A las distancias hacia la izquierda, y hacia abajo, del origen se les da valores negativos, como se señala en la Figura 4. Las cuatro partes en que queda dividido el plano por el eje de las x y por el de las y se llaman *cuadrantes*; tal como se muestra, están indicados por números romanos. En el sistema original de Descartes, un punto tenía que estar en algún lugar del primer cuadrante. En el sistema actual, los puntos correspondientes a diferentes valores de x y y pueden estar situados en cualesquiera de los cuadrantes.

Por supuesto, uno, o ambos, de los dos números de referencia –o coordenadas– de un punto pueden ser negativos. Existen cuatro posibilidades diferentes: (1) Tanto x como y pueden ser positivas; (2) x y y pueden ser ambas negativas; (3) x puede ser positiva y y, negativa; (4) x puede ser negativa y y, positiva. Consideremos algunos ejemplos típicos (Figura 4). Supongamos que las coordenadas son (4, 3); desplazándonos 4 unidades hacia la derecha de O –el origen– sobre el eje de las x, y luego 3 unidades hacia arriba, en forma paralela al de las y, localizamos el punto en A, en el cuadrante I. El segundo par de coordenadas que hemos de considerar (es (−2, 3).



4. Cada uno de los puntos (4,3), (−2, 3), (−4, −2) y (3, −4) está localizado en cuadrantes diferentes.



5. La ecuación de esta recta es $y = 2x$. La coordenada y es aquí siempre el doble de la x.

Vamos dos unidades hacia la izquierda de O a lo largo del eje de la x, y luego 3 hacia arriba. Localizamos el punto en B, en el cuadrante II. Si las coordenadas de un punto son $(-4, -2)$, avanzamos 4 unidades hacia la izquierda de O sobre el eje de las x, y 2 unidades hacia abajo, localizando el punto en C, en el cuadrante III. Finalmente, si las coordenadas son $(3, -4)$, nos corremos 3 unidades hacia la derecha, a partir de O, a lo largo del eje de las x, y luego 4 hacia abajo; el punto está en D, en el cuadrante IV.

El sistema de coordenadas de Descartes, tal como fue extendido por matemáticos ulteriores, es particularmente valioso, puesto que nos permite analizar la figura geométrica descrita por un punto variable. El punto tiene una serie de coordenadas diferentes, a medida que se lo determina para posiciones distintas. Bajo ciertas condiciones, podemos escribir una ecuación que valdrá para todas las posiciones posibles de tal punto; entonces, la ecuación puede emplearse en lugar de la figura geométrica consistente en todos los puntos. Existen ecuaciones para rectas, círculos, elipses, parábolas e hipérbolas.

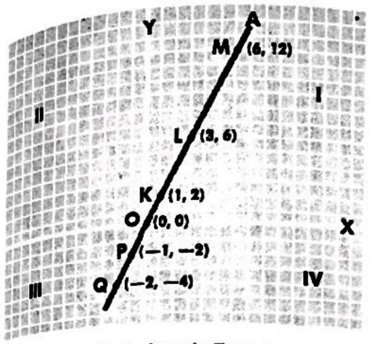
GRÁFICA DE UNA RECTA

Consideremos primero la ecuación para una recta OA, que pasa por el origen O (Figura 5) y luego por los puntos K, L y M. Las coordenadas del punto K son $(1, 2)$. Lo localizamos yendo 1 unidad hacia la derecha de O, y 2 unidades hacia arriba. El punto L tiene las coordenadas $(3, 6)$; es decir, que para localizarlo vamos 3 unidades hacia la derecha de O, 6 hacia arriba. Las coordenadas para M son $(6, 12)$. Se advertirá que, en cada caso, la coordenada y es el doble de la x; por consiguiente, la ecuación que expresa la relación entre x y y para esta línea particular es $y = 2x$.

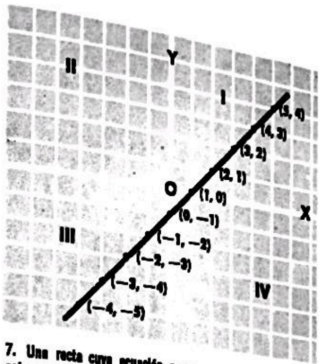
(También podría darse como $x = \frac{y}{2}$).

Todos los valores de coordenadas que hemos presentado hasta el momento para la recta OA son valores positivos, y los puntos cuyas coordenadas son $(1, 2)$, $(3, 6)$ y $(6, 12)$ están todos en el cuadrante I. La recta OA podría extenderse al cuadrante III, tal como muestra la Figura 6. Como ya hemos señalado, la recta pasa por el origen, O. En este punto, las coordenadas son (O, O). Supongamos que la coordenada x es -1. De acuerdo con la ecuación para esta recta, y es siempre igual a $2x$ y, en consecuencia, la coordenada y correspondiente a la x -1 es -2. Localizamos el punto $(-1, -2)$ y continuamos la recta OA hasta este punto (P en el diagrama). Si la coordenada x es -2, la coordenada y debe ser -4. Continuamos la línea hasta el punto O, de coordenadas $(-2, -4)$. La línea podría continuarse indefinidamente tanto en el cuadrante I como en el III. El recorrido descrito por la recta se llama su *gráfica*.

Supongamos que se da la ecuación de una



6. La recta OA, localizada en la Figura 5, se extiende al cuadrante III, localizando los puntos P (−1, −2), y Q (−2, −4).



7. Una recta cuya ecuación es $x - y = 1$. Observar que esta recta se extiende por tres cuadrantes.

recta como $x - y = 1$, y que se nos pide trazar la gráfica de la línea. Primero, estableceríamos una tabla de valores para x y y del modo siguiente:

x	5	4	3	2	1	0	−1	−2	−3	−4	−5
y	4	3	2	1	0	−1	−2	−3	−4	−5	−6

Los valores de y correspondientes a los de x en la tabla pueden deducirse fácilmente, puesto que, en cada caso, el valor y es el valor x menos 1. (Recuérdese que la ecuación de esta recta es $x - y = 1$.) Una vez determinados los valores de x y de y, localizamos los puntos $(5, 4)$, $(4, 3)$, $(3, 2)$, $(2, 1)$, y así sucesivamente, tal como se indica en la Figura 7, y trazamos la línea que conecta los puntos. Ésta es la recta cuya ubicación está indicada por la ecuación $x - y = 1$. Adviértase que pasa por los cuadrantes I, IV y III.

GRÁFICA DE UNA CIRCUNFERENCIA

Podremos deducir la ecuación de una circunferencia si aplicamos el teorema de Pitágoras al sistema de coordenadas cartesianas. Este teorema establece que en un triángulo rectángulo (con un ángulo recto), el cuadrado de la hipotenusa (el lado opuesto al ángulo recto) es igual a la suma de los cuadrados de los otros dos lados (catetos).

Tracemos una circunferencia con centro en el origen y con radio de 5 unidades (Figura 8); cualquiera sea el punto de la circunferencia que consideremos, la distancia x, la distancia y y el radio habrán de formar un triángulo rectángulo, con el radio de 5 unidades como hipotenusa. Según el teorema de Pitágoras, en cada uno de los triángulos de la Figura 8 el cuadrado de la distancia x más el cuadrado de la y será igual al cuadrado de la hipotenusa, es decir, 5^2, o sea 25. En otras palabras, $x^2 + y^2 = 25$, y ésta es la ecuación de la circunferencia que se muestra en la Figura 8.

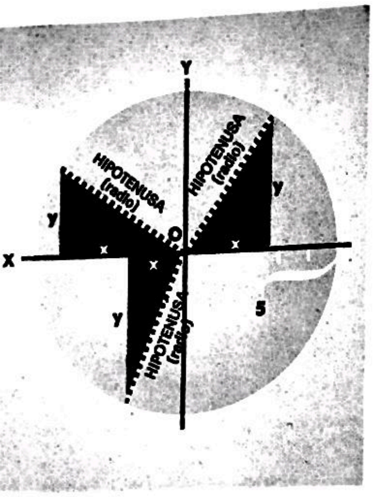
Supongamos que deseamos averiguar si un punto dado pertenece a la circunferencia. Si la suma de los cuadrados de las coordenadas del punto es igual a 25, entonces está en la circunferencia. Sea el punto $(3, 4)$. La suma de 3^2 y 4^2 es 25; luego $(3, 4)$ está en el círculo. Lo mismo rige para $(-4, -3)$, pues la suma de $(-4)^2$ y $(-3)^2$ es 25. Pero el punto $(4, 5)$ no pertenece a la circunferencia, pues la suma de 4^2, ó 16, y 5^2, ó 25, no es igual a 25 (Figura 9).

GRÁFICA DE UNA ELIPSE

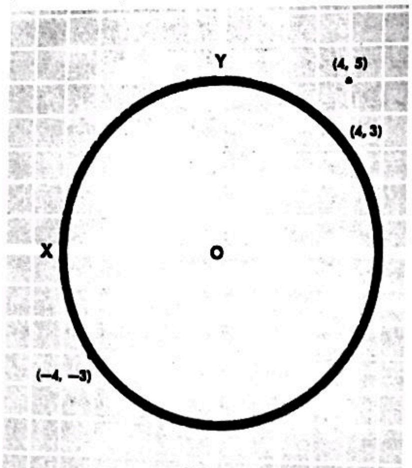
Supongamos que se nos informa que $4x^2 + 9y^2 = 36$ es la ecuación de una figura geométrica. Preparemos una gráfica y veamos qué clase de figura es. Primero debemos establecer una tabla de valores:

x	0	±1	±2	±3
y	±2	±1,9	±1,5	0

El símbolo ± en la tabla significa "más o menos". Por ejemplo, 4 es el cuadrado de +2; lo es asimismo de −2. En consecuencia, la raíz



8. La ecuación de la circunferencia de arriba es $x^2 + y^2 = 25$. El texto explica cómo deducir la ecuación de una circunferencia aplicando el teorema de Pitágoras.



cuadrada de 4 puede ser $+2$ ó -2, o sea, ± 2.

Los valores de la tabla se calculan del siguiente modo: aplicando el valor $x = 0$ a la ecuación $4x^2 + 9y^2 = 36$; $9y^2 = 36$; $y^2 = 4$; $y = \sqrt{4} = \pm 2$. Si x es 1, en la ecuación $4x^2 + 9y^2 = 36$, $4x^2$ es igual a 4. Luego, $4 + 9y^2 = 36$; $9y^2 = 32$; $y^2 = 3{,}6$; $y = \sqrt{3{,}6} = \pm 1{,}9$. Los otros valores de la tabla se calculan de la misma manera.

Localicemos ahora los puntos (0, 2), (0, —2), (1, 1,9), (—1, 1,9), (1, —1,9), (—1, —1,9) y los demás que se dan en la tabla y tracemos una curva continua para unirlos; tal como muestra la Figura 10, la figura geométrica cuya ecuación es $4x^2 + 9y^2 = 36$ resulta ser una elipse.

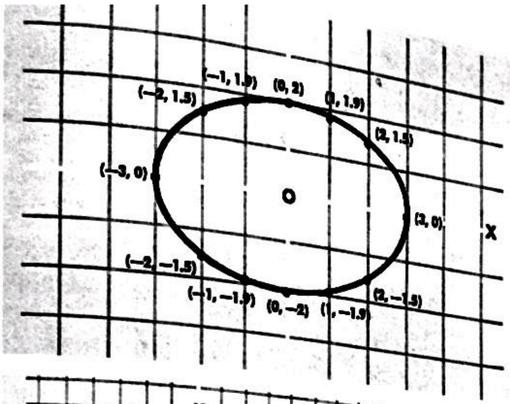
GRÁFICAS DE CURVAS ABIERTAS

Una ecuación típica de una parábola es $x = y^2 - 4$. Preparamos una tabla de valores como sigue:

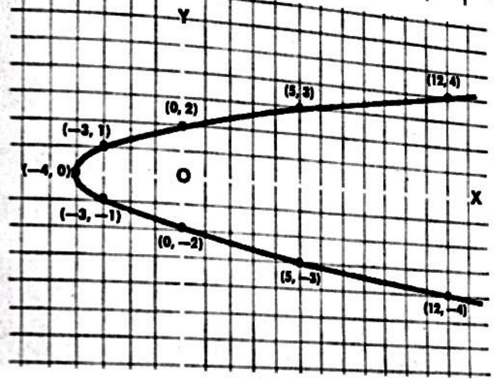
x	—4	—3	0	5	12
y	0	± 1	± 2	± 3	± 4

Si representamos los valores (—4, 0), (—3,

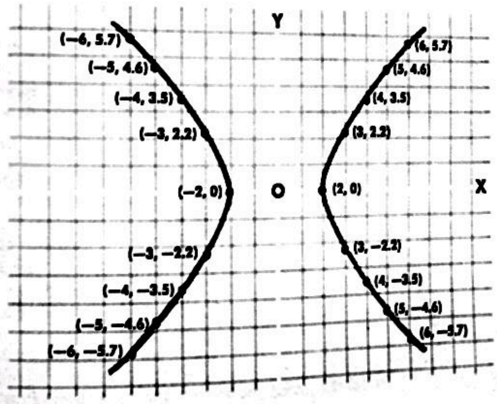
9. La ecuación de la circunferencia de la izquierda es $x^2 + y^2 = 25$. El punto (4,3) está en la circunferencia, puesto que $4^2 + 3^2 = 25$. Lo mismo rige para el punto (–4, –3). Pero el (4,5) no está en esta circunferencia, porque $4^2 + 5^2$ no es igual a 25.



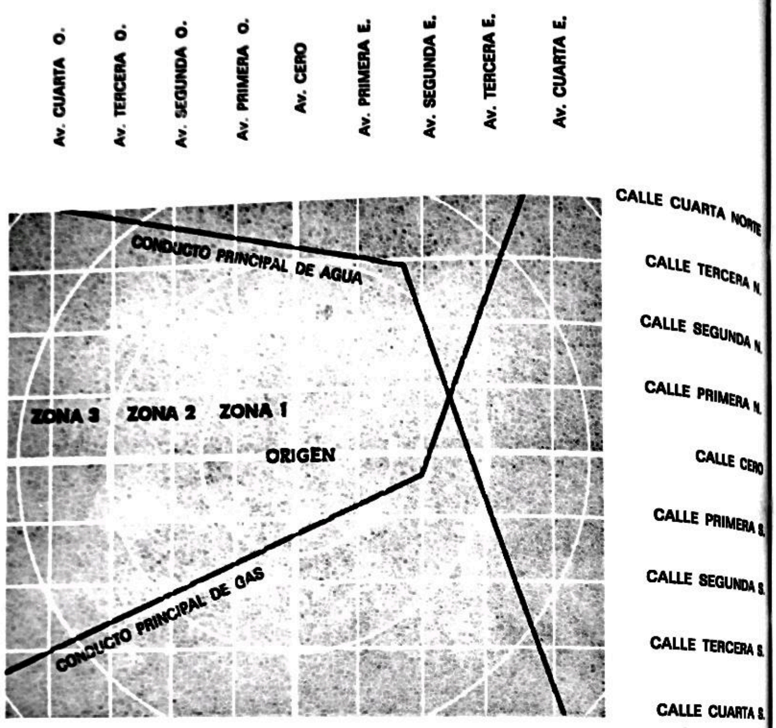
10. Esta figura es una elipse; su ecuación es $4^2 + 9y^2 = 36$.



11. Una parábola con la ecuación $x = y^2 - 4$: es una curva abierta.



12. Una hipérbola ($x^2 - y^2 = 4$). Sus dos ramas son curvas abiertas.



13. A la geometría analítica se le puede dar un uso práctico. Por ejemplo: cañerías maestras de agua y gas pueden localizarse a través de las ecuaciones de rectas, con el origen situado en el centro de la ciudad, tal como se ve en la ilustración de arriba.

1), (−3, −1), (0, 2), (0, −2) y los otros dados en la tabla tenemos la parábola de la Figura 11. Puesto que a mayores valores de x resultan mayores los correspondientes de y, la parábola es una curva abierta que se extiende sin límites.

$x^2 - y^2 = 4$ es la ecuación de una hipérbola típica. Para trazar la gráfica de la figura, primero preparamos una tabla de valores para x y y:

x	± 2	± 3	± 4	± 5	± 6
y	0	$\pm 2{,}2$	$\pm 3{,}5$	$\pm 4{,}6$	$\pm 5{,}7$

Luego, representamos los valores de la tabla, y obtenemos una curva con dos ramas, tal como se muestra en la Figura 12, las que se extienden indefinidamente en ambas direcciones.

USO DE LAS ECUACIONES

Las ecuaciones de las figuras geométricas se pueden sumar y restar, y muchas otras operaciones pueden llevarse a cabo con ellas. Las ecuaciones resultantes de tales operaciones pueden interpretarse por medio de gráficas trazadas sobre la base de valores para x y para y. De este modo, el álgebra puede emplearse para descubrir relaciones geométricas. Como dice el matemático norteamericano Eric T. Bell: "De aquí en adelante, el álgebra y el análisis [es decir la geometría analítica] han de ser nuestros

pilotos en los mares inexplorados del espacio y su geometría".

A la geometría analítica también se le ha dado un uso práctico. Por ejemplo, en muchas ciudades, las cañerías maestras de gas y agua pueden localizarse mediante las ecuaciones de las rectas que emplean el centro de la ciudad como origen (Figura 13). Se pueden usar círculos concéntricos para dar distancias zonales desde el centro de la ciudad. Las intersecciones de círculos y rectas describirán la ubicación aproximada de roturas en las cañerías maestras de agua y gas.

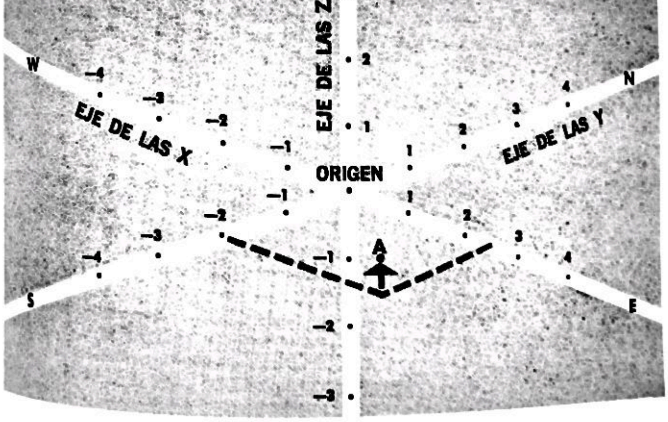
GEOMETRÍA ANALÍTICA DEL ESPACIO

Para localizar un punto en el espacio tridimensional debemos dar la distancia por encima, o por debajo, del plano formado por los ejes de las x y de las y. Agregamos un tercer eje –el eje de las z–, que se extiende rectamente hacia arriba y hacia abajo (Figura 14). Supongamos que deseamos localizar el punto A –que se señala en la figura– en un sistema tridimensional; A está a 2½ kilómetros al este del origen (distancia x); 2 kilómetros al sur del origen (distancia y), y ½ kilómetro hacia arriba (distancia z). Lo localizaríamos dando las coordenadas x, y y z (2,5; 2; 5). En consecuencia, para localizar un punto en geometría analítica del espacio precisamos tres números ordenados: x, y y z como coordenadas. En el espacio tridimensional, si un punto se toma como origen, cualquier otro punto en el espacio puede representarse mediante una terna ordenada de números, y se lo puede localizar en forma definida.

Como en geometría analítica del espacio hay tres dimensiones, habrá tres incógnitas para relacionar en nuestras ecuaciones; así, la ecuación de un plano es $ax + by + cz = k$, donde a, b, c y k son números dados. Si las coordenadas de un punto son tales que satisfacen la ecuación de un plano dado, el punto pertenecerá al plano. Las coordenadas de puntos exteriores al plano no satisfarán la ecuación.

La ecuación de una esfera es $x^2 + y^2 + z^2 = r^2$, siendo r el radio. Es una extensión de la ecuación de la circunferencia de geometría analítica plana: $x^2 + y^2 = r^2$. La ecuación $x^2 + y^2 + z^2 = 25$ representa una esfera cuyo centro está en el origen y con un radio de 5 unidades de longitud. Todos los otros sólidos comunes también pueden representarse por una ecuación, o por una serie de ecuaciones con tres incógnitas: x, y y z. La geometría analítica

14. En la geometría analítica del espacio, tres ejes (x, y y z) corresponden a las tres dimensiones. En el diagrama de abajo, el punto A está 2½ kilómetros al este del origen, 2 kilómetros al sur y ½ kilómetro hacia arriba.



del espacio emplea tales ecuaciones para estudiar los puntos, líneas, superficies y sólidos que existen en el espacio.

GEOMETRIA SUPERIOR

El matemático no se limita al análisis del mundo tridimensional, en el que una terna ordenada de números representa un punto. El matemático pregunta: "¿qué representaría una cuaterna ordenada de números?" Luego, responde diciendo que, por ejemplo (3, 4, 2, 5), localizarían un punto en el espacio tetradimensional. La ecuación $x + y + z + w = 3$ representaría la ecuación de lo que nuestro matemático llamaría hiperplano; (3, 4, 2, 5) sería un punto de este plano. ¿Qué cosa sería hiperplano? El matemático no lo sabe, porque nunca vio uno, empero, esto no lo inhibe para calcular y usar la ecuación de un plano tal, ni para descubrir sus propiedades a través de varias operaciones algebraicas.

¿Cómo puede uno sostener que no existe un mundo de cuatro dimensiones? Supongamos que fuésemos seres bidimensionales que habitan un mundo de dos dimensiones; ignoraríamos que hay un mundo tridimensional, del cual el nuestro, limitado, formaría parte. Sin embargo, este mundo de tres dimensiones sí existe, como bien sabemos. ¿No es perfectamente posible que seres como nosotros, confinados, según creemos, en un mundo de tres dimensiones, realmente estemos viviendo en uno tetradimensional que nuestros sentidos no alcanzan a percibir? También puede ser que este mundo de cuatro dimensiones esté incluido, a su vez, en otro de cinco dimensiones, y éste en uno de seis, y así *ad infinitum*. El matemático encuentra esta especulación fascinante, ya esté o no basada en la realidad.

CIERTAS APLICACIONES PRACTICAS

En verdad, las ecuaciones de estos mundos de más dimensiones poseen ciertas aplicaciones prácticas. Por ejemplo, la ecuación de la hiperesfera –una esfera con más de tres dimensiones– se ha aplicado en la fabricación de tubos de rayos catódicos para televisión.

Vimos que la ecuación de la esfera tridimensional, con la que estamos familiarizados, es $x^2 + y^2 + z^2 = r^2$, con r^2 que representa el cuadrado del radio. Por analogía con esta ecuación, el matemático define una esfera en cuatro dimensiones, como $x^2 + y^2 + z^2 + w^2 = r^2$, donde x, y, z y w son las cuatro dimensiones. Ésta es la primera hiperesfera, y no hacemos esfuerzo alguno por dibujarla. Continuamos luego con las ecuaciones de una esfera pentadimensional, una hexadimensional, y así sucesivamente. Podemos definir una hiperesfera de n dimensiones, como

$$x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 = r^2$$

donde x_1, x_2, x_3, y así sucesivamente hasta x_n, representan las diferentes dimensiones.

Veamos ahora cómo puede aplicarse esta ecuación a la fabricación de tubos para televisión. El fabricante de los mismos querría que todos tengan exactamente la misma longitud (40 centímetros, digamos). Sin embargo, como los tubos nunca son de la misma longitud exactamente, permitirá una desviación de no más de 0,012 centímetros; si la desviación es mayor, el tubo no será satisfactorio.

Se selecciona un muestrario de 10 tubos con fines de ensayo, extraídos de una tanda de 1.000; se mide la longitud de cada uno, y se anota la desviación con respecto a 40 centímetros. La desviación de cada tubo debe constituir una dimensión; por consiguiente, tenemos 10 dimensiones en total. Elevamos al cuadrado cada una de las desviaciones, y luego las sumamos, según la ecuación de la hiperesfera que se dio antes:

$$x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2 = r^2$$

Para una desviación permisible de no más 0,012 centímetros, r^2 no debe exceder de cierta cantidad. Si tomamos un muestrario distinto de otros 10 tubos, esperaríamos una suma diferente cuando hacemos la adición de los cuadrados de sus desviaciones con respecto a 40 centímetros; esto es, un valor diferente de r^2. El estadígrafo estudia las variaciones del valor r^2 de un muestrario a otro. De ahí en adelante, si el fabricante desea probar una tanda de 1.000 tubos, toma una muestra de 10. Si el valor r^2 de ésta es suficientemente pequeño, se acepta la tanda completa de 1.000 tubos, por satisfacer las especificaciones requeridas. Así, la hiperesfera de los matemáticos, que al principio sólo parecía ser especulación nada práctica, tiene una aplicación determinada: la de ayudar a comprobar la calidad de la recepción de nuestro televisor.

GEOMETRÍA NO EUCLIDIANA

La geometría que se presenta en los artículos Geometría Plana o Geometría del Espacio, y que se enseña en los colegios secundarios, se denomina geometría euclidiana. Se denomina así porque está basada en el sistema del matemático griego Euclides, por él enseñado en Alejandría, Egipto, alrededor del 300 a.C. Estableció una serie lógica de teoremas, o enunciados ordenados de forma tal que la demostración de cada uno depende de (1) los teoremas precedentes y (2) de ciertas suposiciones, o postulados. Euclides llamó a estos supuestos nociones comunes y los aceptó sin demostración.

A fin de probar algunos de los teoremas de la primera parte de su sistema, Euclides halló necesario suponer que por un punto exterior a una recta dada sólo podía trazarse una paralela a aquélla (Figura 1). Intentó afanosamente probar esto, pero fracasó. Finalmente, tuvo que considerar el enunciado sobre paralelas (o mejor, un enunciado equivalente) como noción común o postulado, que debía aceptarse por autoevidente. Todo el sistema de Euclides depende de la validez de esta noción común particular, que ha sido llamada *postulado de Euclides*.

En los siglos siguientes, muchos matemáticos trataron de probar que, por un punto exterior a una recta dada, sólo podría trazarse una paralela a aquélla; no tuvieron más éxito que Euclides. Debieron aceptar el postulado de Euclides como incontrovertible, pero no demostrado.

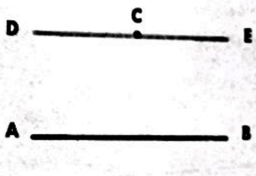
En épocas modernas, no todos los matemáticos concedieron que el postulado fuese axiomático; se creyeron justificados para hacer distintos supuestos y construyeron geometrías basadas en éstos. El matemático alemán del siglo XIX, G. F. Bernhard Riemann, basó su geometría –llamada riemanniana– en el postulado de que dos líneas nunca son paralelas. También en ese siglo, el ruso Nikolai Ivanivich Lobachevsky e, independientemente, el húngaro Janos Bolyai, crearon otra geometría. Se basaba en el supuesto de que, por un punto dado, pueden trazarse, por lo menos, dos paralelas a una línea dada. Por ello, a las geometrías de Riemann y Lobachevsky-Bolyai se las conoce como *no euclidianas*.

SUPUESTOS DIFERENTES

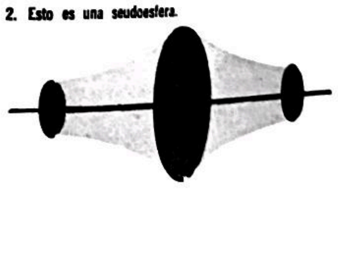
Para comprender la diferencia entre las geometrías euclidianas y no euclidianas, consideremos la distancia más corta entre dos puntos en (1) un plano, (2) una esfera y (3) una seudoesfera, que se muestra en la Figura 2 y sugiere dos cestos de papeles unidos por la boca. La distancia más corta entre dos puntos sobre cualquier superficie se denomina línea *geodésica*.

En un plano, hemisfera o hemiseudoesfera (media seudoesfera) medimos la misma distancia *AB* usando una geodésica (la distancia más corta posible) en cada caso (Figura 3). En un plano –tal como el de la Figura 3a– la geodésica es una recta. En una hemiesfera (Figura 3*b*) es una arco de círculo máximo (el mayor círculo que puede trazarse sobre una

1. Según la geometría euclidiana, a través del punto C sólo puede trazarse una, y sólo una, recta DE paralela a la AB.



2. Esto es una seudoesfera.



esfera). En la hemiseudoesfera (Figura 3*c*), *AB* está en el círculo que marca la "cintura" de la seudoesfera.

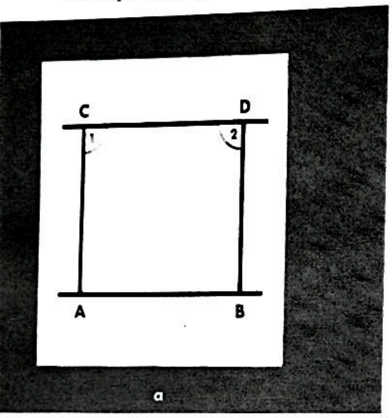
En el plano, hemiesfera y hemiseudoestera, trazamos la geodésica *AC* perpendicular a *AB* en *A*, y la geodésica *BD*, longitud igual a *AC* y perpendicular a *AB* en *B*. Finalmente, trazamos la geodésica *CD*, uniendo *C* y *D*. En el plano, *CD* es una recta; en la hemiesfera, un arco, *CxD*, que forma parte del círculo máximo *MCDN*; en la hemiseudoesfera, es el arco *CyD*.

¿Cuán grandes son los ángulos *1* y *2* en las Figuras 3*a*, 3*b* y 3*c*? Las respuestas a estas preguntas ilustrarán las diferencias entre la geometría euclidiana y las no euclidianas de Riemann y Lobachevsky-Bolyai.

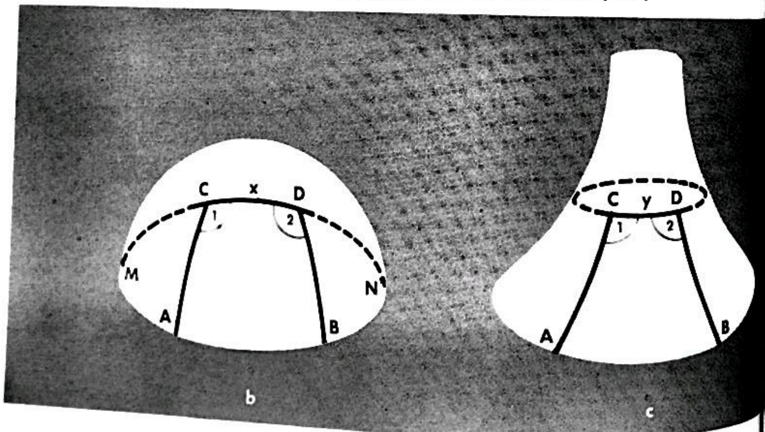
En el caso del plano de la Figura 3*a*, parecería obvio que los ángulos *1* y *2* son rectos (de 90°); pero no podemos probar esto como teorema geométrico, a menos que coincidamos, primero, en que por *C* pasa sólo una línea, la *CD* paralela a la *AB*. Si partimos de ese supuesto, estamos aceptando el postulado de Euclides.

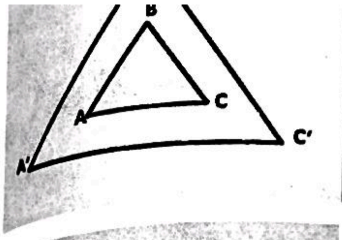
En la hemiesfera (Figura 3*b*), aparentemente, los ángulos *1* y *2* son obtusos –mayores de 90 °–. ¿Podemos probarlo? No, a menos que supongamos primero que cada geodésica que pase por *C* cortará la línea *AB* en dos puntos. Si admitimos esto, podemos probar que los ángulos *1* y *2* son mayores de 90°. Riemann supuso que, inclusive en un plano, cualquier línea (tal como *CD* de la Figura 3*a*) trazada por un punto exterior (tal como *C*) se encontrará con cualquier otra (*AB*) en dos puntos; por eso, en su geometría no existen líneas paralelas. Sobre esta suposición, desarrolló un conjunto de teoremas perfectamente lógico.

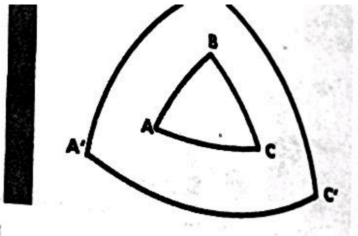
"Pero cualquiera puede ver que la línea *CD* de la Figura 3*a* es paralela a *AB*, y que las dos nunca se encontrarán", argüirán ustedes. Mas,

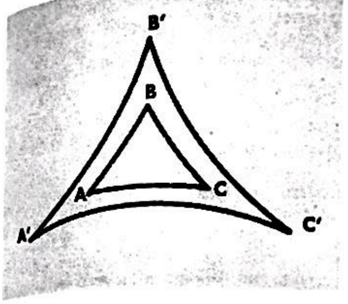


3. En *a*, *ABCD* se ha trazado en un plano. En *b*, se lo ha hecho sobre una hemiesfera; en *c*, sobre una hemiseudoesfera (mitad de una seudoesfera). En cada caso, *AC* y *BD* son iguales y perpendiculares a la línea *AB*. Todas las líneas que aquí se muestran son geodesicas; es decir, representan las distancias más cortas entre los puntos.









6. Arriba: en la geometría de Lobachevsky-Bolyai, la suma de los ángulos de un triángulo siempre es menor que 180°. Disminuye a medida que aumenta el área del triángulo.

en tanto no se pruebe esto, no puede negarse que es perfectamente lógico desarrollar una nueva clase de geometría basada en otro supuesto.

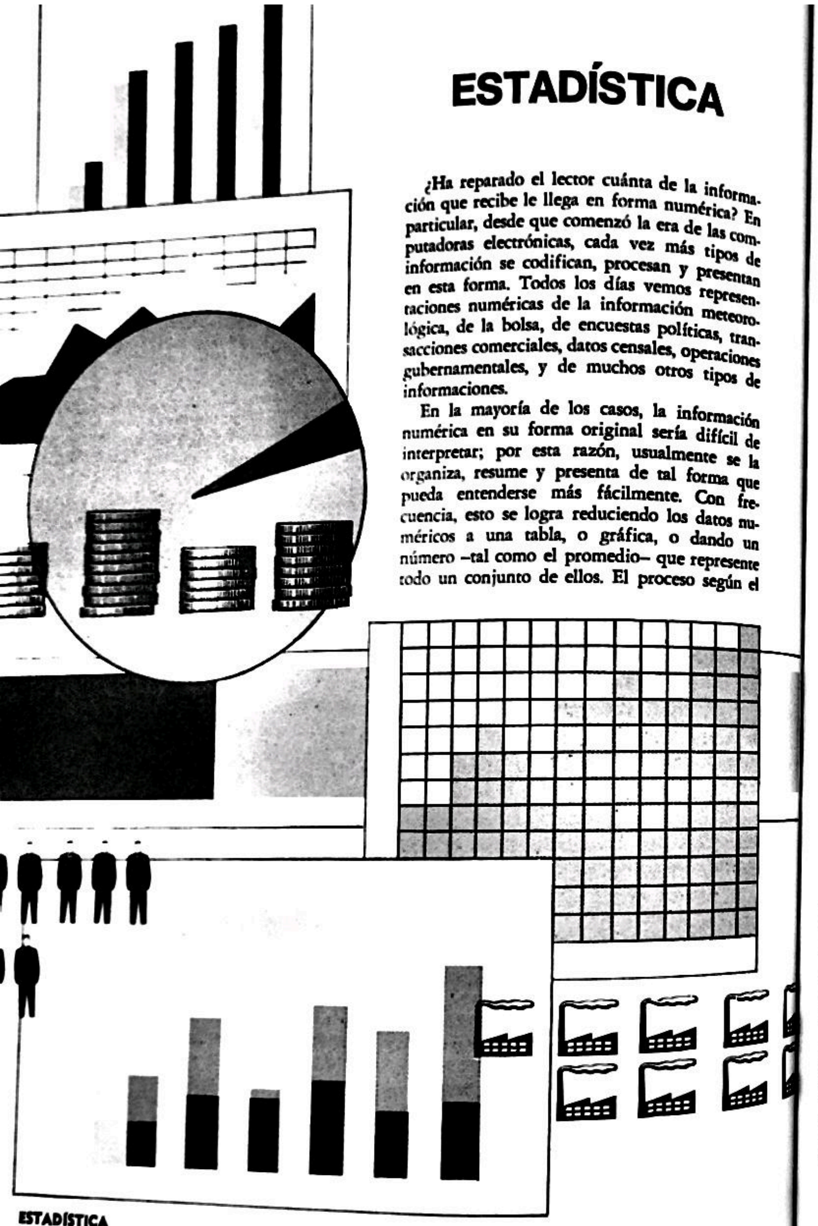
Examinemos ahora los ángulos *1* y *2* de la hemiseudoesfera (Figura 3*c*); aparentemente son agudos (de menos de 90°). Tampoco podemos probar esto, a menos que supongamos, en este caso, que existen dos geodésicas que pasan por *C* que nunca cortarán *AB*. Lobachevsky y Bolyai lo hicieron y lo aplicaron a todo tipo de superficie, incluyendo las de esferas y planos. Sus geometrías describen un mundo completamente diferente al del Euclides y Riemann; un mundo en el que, por cada punto exterior, pasan dos paralelas a una línea dada.

4. Arriba a la izquierda: en la geometría de Euclides, la suma de los ángulos de un triángulo, tal como el *ABC*, es 180°. Si se aumenta el área de tal triángulo –de modo que tengamos el triángulo *A'B'C'*–, la suma de los ángulos sigue siendo igual a 180°.

5. Arriba a la derecha: según la geometría de Riemann, la suma de los ángulos de un triángulo, tal como el *ABC*, siempre es mayor que 180°. Aumenta a medida que lo hace el área del triángulo, tal como en el triángulo grande de *A'B'C'*.

Como es obvio, tres geometrías basadas en supuestos tan diferentes pueden exhibir muchas diferencias notables. Considérese, p. ej., la suma de los ángulos de cualquier triángulo. En la geometría euclidiana es de 180°. Se conserva así pese al aumento de tamaño de aquél (Figura 4). En la geometría riemanniana, dicha suma es siempre mayor que 180°. A medida que aumenta el área del triángulo, lo hace la suma de sus ángulos (Figura 5). En la geometría de Lobachevsky-Bolyai, tal suma es siempre menor que 180°, y disminuye a medida que aumenta el área del triángulo (Figura 6).

No puede decirse que una de estas geometrías sea correcta y que las otras no lo sean: más bien son distintas explicaciones del espacio, basadas en supuestos distintos. Aceptamos la que explica más satisfactoriamente un fenómeno particular que nos interese. Por ejemplo, la geometría riemanniana ha explicado mejor ciertos aspectos astronómicos de la teoría de la relatividad de Einstein que las geometrías euclidiana o de Lobachevsky-Bolyai.



ESTADÍSTICA

¿Ha reparado el lector cuánta de la información que recibe le llega en forma numérica? En particular, desde que comenzó la era de las computadoras electrónicas, cada vez más tipos de información se codifican, procesan y presentan en esta forma. Todos los días vemos representaciones numéricas de la información meteorológica, de la bolsa, de encuestas políticas, transacciones comerciales, datos censales, operaciones gubernamentales, y de muchos otros tipos de informaciones.

En la mayoría de los casos, la información numérica en su forma original sería difícil de interpretar; por esta razón, usualmente se la organiza, resume y presenta de tal forma que pueda entenderse más fácilmente. Con frecuencia, esto se logra reduciendo los datos numéricos a una tabla, o gráfica, o dando un número –tal como el promedio– que represente todo un conjunto de ellos. El proceso según el cual se recogen y presentan finalmente datos numéricos en forma utilizable y comprensible constituye una parte importante de la matemática estadística.

La estadística no sólo es muy útil para la comunicación sino que también brinda una base para tomar decisiones. El gobierno hace uso extensivo de ella cuando estima sus necesidades presupuestarias y establece las tasas impositivas. La estadística permite a los fabricantes comparar procesos de elaboración cuando buscan mejorar sus productos o aumentar sus ganancias. Los gerentes de almacenes pueden confiar en los análisis estadísticos para determinar qué artículos deben acopiar. Los científicos la emplean para comparar los efectos de variables críticas en sus experimentos. Las compañías de seguros triunfan o fracasan según la precisión de sus predicciones estadísticas. Los ingenieros basan el diseño de carreteras y puentes en los estudios estadísticos de materiales y tráfico. Los funcionarios docentes pueden modificar sus programas basándose en análisis estadísticos de necesidades y logros de los alumnos; la lista de usos de la estadística es prácticamente interminable.

RECOPILACIÓN DE DATOS

La ciencia de la estadística abarca variedad de tareas. Inclusive, el trabajo aparentemente simple de recoger datos numéricos requiere un estudio cuidadoso. Evidentemente, las conclusiones de un estudio estadístico no pueden ser más confiables que las cifras sobre las que se basan. El estadígrafo debe estar seguro de que los datos reunidos son precisos y atinentes al problema que se estudia, así como representativos del mismo. Conclusiones no válidas extraídas de la evidencia estadística se deben a menudo a insuficiencias de los datos recogidos. Si bien no podemos tratar en detalle el asunto de la recopilación de datos, su importancia es tan grande que, por lo menos, debemos tener conocimiento de algunos de los problemas implicados.

Primero: La *población* que debe estudiarse tiene que estar muy bien definida. ¿Qué significamos al decir "población"? Para el estadígrafo, puede consistir en un conjunto de ciudades, automóviles, libros o, inclusive, experimentos científicos. En verdad, la población de un estudio estadístico podría ser cualquier conjunto de objetos con una característica común a la que pudiera asignársele un número. Por supuesto, la población seleccionada debe suministrar los datos numéricos apropiados al problema que se estudia. Si la población para un estudio estadístico no está bien definida o si no es representativa del problema, los resultados de aquél serán difíciles de interpretar o aplicar. Por ejemplo, las encuestas sobre preferencias electorales de estudiantes secundarios tendrían un valor cuestionable para predecir los resultados de una elección, puesto que, como se sabe, pocos estudiantes de secundaria pueden votar.

Una vez identificada la población, las características particulares que hay que estudiar deben representarse en forma numérica. A veces, los datos numéricos ya se consiguen en forma registrada; por ejemplo, quien desee estudiar la precipitación pluvial en su ciudad durante el año anterior, probablemente podría obtener los datos necesarios del servicio meteorológico local; en este caso, la población podría definirse como el conjunto de días en el año.

A veces, los datos numéricos pueden obtenerse mediante un simple proceso de recuento. Para alguien interesado, por ejemplo, en un estudio de los libros de la biblioteca de un colegio, este proyecto implicaría contar el número de volúmenes dedicados a cada una de varias materias.

Con más frecuencia, los datos para un estudio estadístico se obtienen midiendo alguna característica común de los hechos estudiados. Si éstos fueran un conjunto de experimentos científicos, al investigador podrían interesarle características tales como tiempo, temperatura, volumen y masa: en cada caso, necesitaría emplear un instrumento de medición apropiado para asignar un número a la característica.

En algunas ocasiones no existe instrumento de medición, y el investigador debe crearlo. Con este fin, puede proyectar una indagación que luego se convertirá en el instrumento para asignar un número que mida los progresos de los estudiantes. Al igual que para cualquier otro instrumento de medición, la precisión es consideración primordial. El investigador necesitaría saber hasta dónde el número asignado representa el verdadero valor de la característica que se mide; en este caso, el adelanto de los alumnos.

TABLA I

Treinta y cinco calificaciones de una prueba de ortografía

16	18	16	14	12	13	18
19	12	17	15	16	14	16
14	16	20	16	17	15	15
16	15	14	18	15	16	13
17	13	16	17	15	19	17

En última instancia, el valor de un estudio estadístico depende, en gran medida, de la calidad del instrumento de medición que se emplea. Por esta razón, la construcción y evaluación de dicho instrumento dada su importancia, es a menudo tarea delicada de un estudio estadístico.

Algunas veces, es posible obtener datos numéricos sobre cada miembro de la población que se estudia. Cuando sucede esto, los datos son completamente representativos de ésta, y la tarea del estadígrafo es describir los datos numéricos obtenidos; a este ramo de la estadística se lo llama ***estadística descriptiva***.

EXTRACCIÓN DE MUESTRAS DE UNA POBLACIÓN

A menudo es necesario, o práctico, recoger datos sólo para una *muestra* de la población y hacer *inferencias estadísticas* sobre dicha población. Inferir es deducir una cosa desconocida de otra conocida, y una inferencia estadística es la que se basa en datos estadísticos. Cuando sólo existen datos disponibles para una muestra, ésta representa lo conocido, y la población, lo desconocido. Cualquier subconjunto de una población constituiría una muestra, pero las inferencias estadísticas sólo son válidas cuando la muestra es representativa de la población. Los estadígrafos emplean muchas técnicas para asegurarse de que las muestras seleccionadas sean representativas.

Cuando cada miembro de la población tiene igual probabilidad de ser elegido, tenemos lo que se denomina una *muestra al azar* o *aleatoria*; se da por sentado usualmente que es representativa de la población. En algunos problemas especiales, el estadígrafo emplea una *muestra estratificada*, la cual asegura que en la muestra están representadas secciones específicas de la población en estudio. El proceso de identificación de una muestra representativa es una tarea clave de muchos estudios estadísticos. En los ejemplos que siguen, supondremos que las muestras empleadas son representativas de las poblaciones de cuyo seno se han seleccionado.

Permítaseme resumir las observaciones anteriores sobre recopilación de datos con el ejemplo siguiente: supongamos que la población que se estudia es el conjunto de alumnos de un grado determinado. Si cada uno de éstos fuera asignado a una clase de inglés según un proceso aleatorio, la clase de inglés constituiría una muestra representativa de la población. Si se midiera la altura, peso o edad de cada estudiante de la clase, o registráramos su calificación en una prueba particular, o si contáramos el número de personas de su familia, obtendríamos un conjunto de números que serían, entonces, datos para un estudio estadístico, y podríamos emplearlos para describir la clase de inglés (la muestra). También podríamos usarlos para hacer *estimaciones*, o *inferencias*, sobre el conjunto total de estudiantes del grado (la población).

ORGANIZACIÓN DE LOS DATOS

Una vez recogidos, los datos deben disponerse en algún orden sistemático antes de que pueda hacerse una interpretación útil, o extraer conclusiones. A veces, una simple tabla o gráfica puede ser apropiada como primer paso hacia el análisis estadístico de datos numéricos.

Los que se presentan en la Tabla I son calificaciones obtenidas por treinta y cinco alumnos en una prueba de ortografía. Cada una representa el número de palabras escritas correctamente por un estudiante dado. En este caso, la palabra calificación se emplea en su sentido usual. Empero, independientemente de la naturaleza de los datos numéricos, los estadígrafos usan con frecuencia el término "calificación bruta" para indicar los números individuales obtenidos como base para un estudio estadístico.

Es difícil hacer una interpretación útil de los datos de la Tabla I. El modo más simple de organizarlos sería disponerlos en orden numérico. Es común registrar solamente las diferentes calificaciones tomadas (en este caso, 16, 18, 14 y así sucesivamente) y anotar la *frecuencia* con que aparece cada una de ellas. La Tabla II da la frecuencia con que se presentan los mismos datos que aparecen en la Tabla I. Incluso un ligero examen de la Tabla II permite alguna interpretación elemental de los datos. Podemos observar fácilmente las calificaciones más altas y más bajas (20 y 12) y la más frecuente (16). Inclusive, podemos comenzar a tener una idea de lo que buscamos por el modo en el que parecen agruparse en torno de un punto central: en este caso, la calificación 16.

Podemos obtener una aclaración ulterior de los datos si damos a la Tabla II una forma gráfica. La Figura 1 es un tipo de gráfica común que se usa para presentar distribuciones de frecuencia. Los números por debajo de la línea horizontal representan calificaciones; los de la izquierda, la frecuencia de distribución, es decir, el número de veces que aparece cada una. Ésta es una gráfica de frecuencias.

El polígono de frecuencias que se muestra en la Figura 2 se basa en la misma idea. En este caso, podemos considerar que se han trazado líneas perpendiculares a las calificaciones de la línea inferior y a las frecuencias de la izquierda. Un punto indica la intersección de una línea de calificación y una de frecuencia. Así, tenemos un punto en el que se encuentran la calificación 12 y el número de frecuencia 2, y otro en el que lo hacen la calificación 16 y el número de frecuencia 9. Los puntos están unidos mediante rectas. En cada caso, se traza una línea entre el punto que representa la calificación más baja y la precedente en la línea horizontal; también entre el punto que representa la más alta y la siguiente en la línea horizontal.

Un examen de estas dos gráficas revelará exactamente la misma información disponible en la Tabla II. Para algunas personas, la forma gráfica es más fácil de interpretar que la tabular; en particular, la tendencia de las califi-

TABLA II

Distribución de frecuencias de calificaciones de ortografía

Calificación	Frecuencia
20	1
19	2
18	3
17	5
16	9
15	6
14	4
13	3
12	2
	n = 35

Nota: "n" significa número de calificaciones.

TABLA III

Tablas de frecuencias agrupadas de 200 calificaciones de ortografía

Intervalo	Frecuencia
23-25	8
20-22	19
17-19	44
14-16	60
11-13	40
8-10	21
5- 7	8
	n = 200

caciones a agruparse en torno de un punto central se vuelve más evidente cuando los datos se representan en forma gráfica.

AGRUPAMIENTO DE CALIFICACIONES BRUTAS

En muchos estudios estadísticos, el número de calificaciones brutas diferentes obtenidas es tan grande que se vuelve engorroso trabajar con tal cantidad de valores distintos. En estos casos, es común condensar estos datos, agrupando las calificaciones brutas en *intervalos de clase*; en vez de considerar cada una en forma individual, combinamos cierto número de las que sean adyacentes para formar un intervalo. Así, combinaríamos las individuales 23, 24 y 25, para formar el intervalo 23-25. Los intervalos se tratan en forma muy parecida a como trataríamos calificaciones individuales.

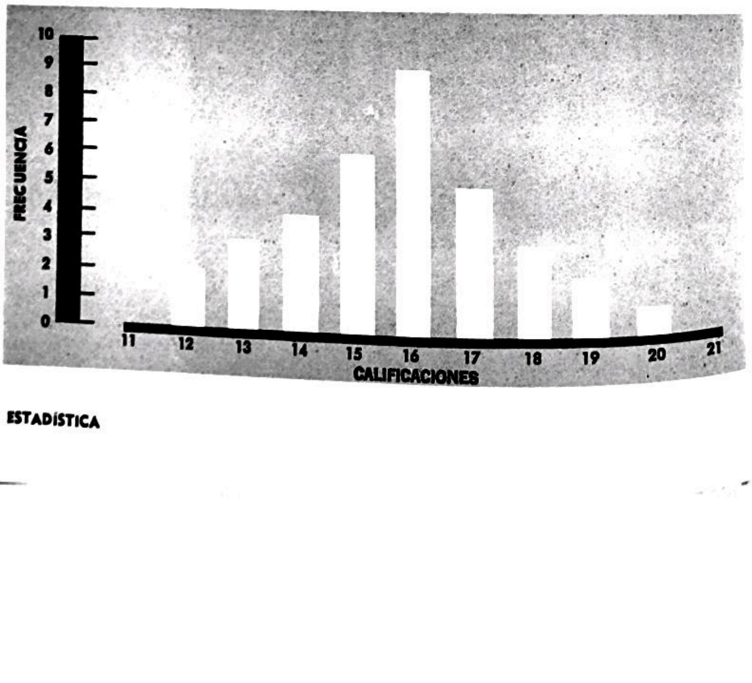
En la Tabla III se muestra una tabla de frecuencias agrupadas, basada en intervalos; presenta calificaciones obtenidas por 200 estudiantes en la misma prueba de ortografía a que se refiere la Tabla I. Podríamos suponer que las 35 calificaciones de la antedicha Tabla son representativas de las calificaciones de una de las seis clases de inglés asignadas aleatoriamente, cuyos resultados totales se dan en la Tabla III.

Cuando el número de calificaciones brutas posibles es grande, el agrupamiento de datos en intervalos apropiados permite al investigador trabajar con un número manejable. Los siete intervalos de la Tabla III pueden reflejar hasta tres veces el número de calificaciones brutas diferentes; al mismo tiempo, este método de agrupamiento tiene la desventaja de ocultar las individuales. Así, en el intervalo 23-25 sólo puede representarse una mientras las tres podrían serlo en otro intervalo. Cuando trabajamos con intervalos de clase debemos olvidar esta consideración. Una vez condensados los datos por agrupamiento, en todos los análisis y cómputos siguientes debemos tratar las calificaciones individuales como si estuvieran distribuidas uniformemente en todo el intervalo al que pertenecen.

Las representaciones gráficas de tablas de frecuencias agrupadas son muy similares a las que se presentaron antes para datos sin agrupar. Es digna de mención una clase especial de gráfica, el *histograma de frecuencia*. Para construir un histograma a partir de la Tabla III –tal como en la Figura 3– la frecuencia de calificaciones en cada intervalo se representa por un rectángulo con centro en el punto medio del intervalo, altura igual a la frecuencia y ancho igual al del intervalo. Debemos establecer aquí una diferencia.

Es necesario distinguir entre límites de *calificación* y límites *reales:* cuando un intervalo se identifica como 14-16, los límites referidos son los *límites de calificación*, e identifican a la más baja y a la más alta pertenecientes al mismo intervalo. A efectos de tratamiento matemático y representación gráfica, es común el uso de los límites reales 13,5-16,5 para identificar este mismo intervalo; se lo representa

1. Gráfica de frecuencia de los datos de la Tabla II.



2. A menudo, las gráficas son más fáciles de interpretar que las tablas.

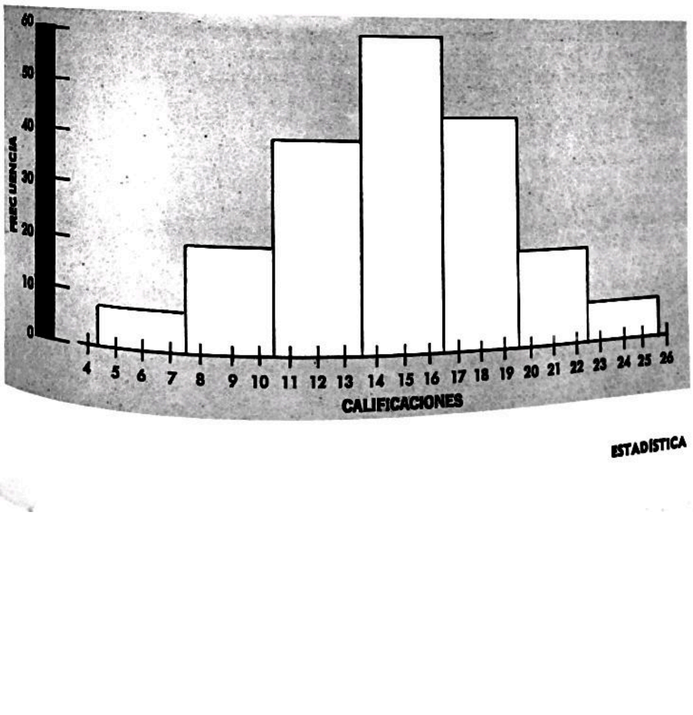
como extendiéndose a mitad de camino entre las calificaciones que lo preceden y las que siguen inmediatamente. Tal interpretación es consecuente con el modo en que generalmente indicamos las mediciones. Por ejemplo, si estuviéramos midiendo con aproximación al metro, cualquier medición entre 13 metros, 50 centímetros y 14 metros, 50 centímetros, se consideraría como 14 metros.

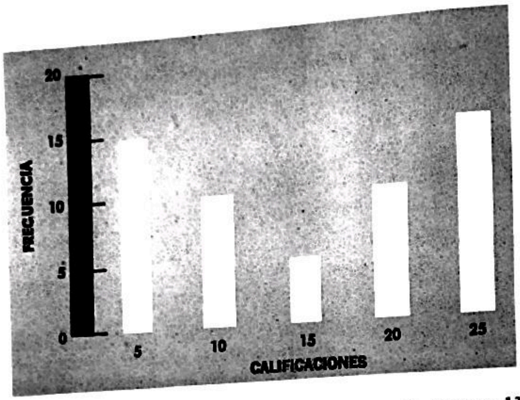
ANÁLISIS DE LOS DATOS

Las tablas y gráficas pueden ayudarnos bastante a comprender un conjunto de calificaciones. Sin embargo, para muchos fines, es preferible tratar de representar un conjunto completo de éstas mediante una sola. Cuando se selecciona un número único para representar todo un conjunto, en lo primero que se piensa usualmente es en el promedio. Tal como señalamos antes, las calificaciones que hemos estado examinando parecen agruparse en torno de un punto central; es este punto de centralización el que los estadígrafos identifican cuando se refieren a una calificación promedio. En estadística existen varias clases de promedios. Tres son comunes en el análisis estadístico: moda, mediana y media. A cada una se la llama medida de centralización. Para los datos con los que hemos estado trabajando en este artículo, moda, mediana y media están próximas, pero éste no es siempre el caso; pueden ser bien diferentes.

La *moda* se identifica muy fácilmente a partir de una tabla, o gráfica, de frecuencias. Es la calificación que aparece más a menudo; en un sentido, la más popular. En los datos no agrupados que se presentaron arriba ya hemos observado que la más frecuente es 16. Esto puede determinarse a partir de la Tabla II, Figura 1 ó Figura 2. Así, 16 es la moda de este conjunto de calificaciones. En los datos agrupados, la *moda bruta* se identificaría como el punto medio del intervalo con la frecuencia más alta. Ya sea en la Tabla III, o en la Figura 3, podemos ver que el intervalo de frecuencia más

3. Histograma de frecuencia de datos de la Tabla III.





4. En la gráfica de esta distribución hay dos calificaciones –5 y 25– que tienen la frecuencia mayor. Compárese esta gráfica con la de la Figura 5.

elevada es 14-16; en consecuencia, 15 sería la moda bruta de esta distribución, puesto que es el punto medio del intervalo. Como las calificaciones individuales no se mencionan no podemos estar seguros de que sea en realidad la individual la más frecuente, pero constituye nuestra mejor estimación de ésta.

La *mediana* es la calificación media en un conjunto de calificaciones. Si examinamos el de la Tabla II, vemos que la mediana es 16. Si hubiese habido un número par de calificaciones, la mediana se habría dado como una cifra intermedia entre las dos medias.

Como el segundo conjunto de datos, presentados en la Tabla III y la Figura 3, ha sido condensado por agrupamiento, no podemos trabajar con valores individuales y debemos hallar un nuevo procedimiento para identificar la mediana. Según la Tabla III, encontramos que el puntaje medio debe caer en el intervalo 14-16; como sólo 69 calificaciones del total de 200 caen por debajo de él, sabemos que 31 de las 60 del intervalo 14-16 también deben estar por debajo de la mediana. Puesto que debemos suponer que, para los datos agrupados, las calificaciones se hallan uniformemente distribuidas dentro de un intervalo, supondremos que la mediana está a $^{31}/_{60}$ del ancho del intervalo, por encima de su límite inferior. $^{31}/_{60} \times 3$ (número de calificaciones en el intervalo) $=$ $^{93}/_{60} = 1{,}5 +$. Sumamos este número al punto extremo izquierdo del intervalo 14-16. Señalamos anteriormente que este punto es 13,5 ya que un intervalo comienza y termina a mitad de camino entre dos calificaciones; en consecuencia, tenemos $13{,}5 + 1{,}5 = 15$. Luego, la mediana es 15.

La *media* es la medida de centralización más empleada comúnmente, y es el promedio en el que todos pensamos primero. Se la encuentra dividiendo la suma de todas las calificaciones individuales por el número de ellas en el conjunto. Este cálculo puede abreviarse cuando se tiene a mano una tabla de frecuencias, multiplicando cada calificación por su frecuencia, y haciendo luego la suma. Por ejemplo, en la Tabla II efectuaríamos el cálculo siguiente: $(20 \cdot 1) + (19 \cdot 2) + (18 \cdot 3) + (17 \cdot 5) + (16 \cdot 9) + (15 \cdot 6) + (14 \cdot 4) + (13 \cdot 3) + (12 \cdot 2)$, dividido por $35 = 550$, dividido por $35 = 15{,}7$; 15,7 es la media.

Cuando se computa la media para datos agrupados suponemos que las calificaciones de cualquier intervalo están distribuidas uniformemente. Multiplicamos el valor del punto medio de cada intervalo por la frecuencia, y dividimos la suma de los números resultantes por el total de calificaciones. En el caso de los datos agrupados de la Tabla III, podríamos computar la media así: $(24 \cdot 8) + (21 \cdot 19) + (18 \cdot 44) + (15 \cdot 60) + (12 \cdot 40) + (9 \cdot 21) + (6 \cdot 8)$ dividido por $200 = 3.000$, dividido por $200 = 15$.

El cálculo de la media puede simplificarse si se usan métodos más avanzados, que son matemáticamente equivalentes a los cálculos anteriores.

La moda se usa menos frecuentemente que la mediana o la media. Es útil sólo cuando queremos identificar el número que aparece más repetidamente en un conjunto. A decir verdad, si la moda ha de ser verdaderamente significativa, un número de un conjunto debe aparecer con frecuencia considerablemente mayor que cualquier otro. La ventaja de la moda es que, al igual que la mediana, es más fácil de identificar y entender; pero el término "moda" es ambiguo a veces, pues puede existir más de una calificación con la "frecuencia más alta" (ver Figura 4). Asimismo, la moda no es muy confiable como indicación de centralización, porque la calificación más común no siempre está cerca del centro de una distribución dada.

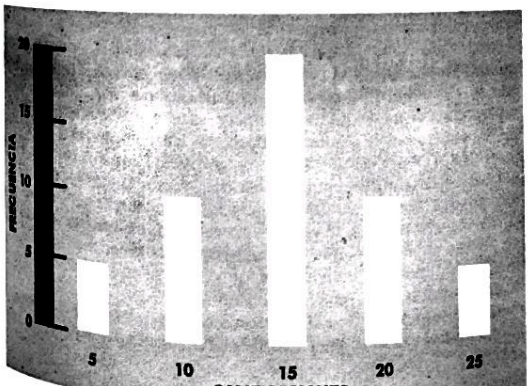
La ventaja principal de la mediana es no ser afectada por cifras extremas. El ingreso "promedio" de una comunidad, por ejemplo, es a menudo reflejado con más precisión por la mediana que por la media, porque el valor de aquélla no es influido por unos pocos ingresos muy elevados o muy bajos. La idea de mediana está íntimamente relacionada con el concepto de percentiles: un tipo de calificación que reciben los estudiantes en ciertas pruebas clásicas del colegio. La mediana corresponde al 50avo percentil. También pueden emplearse otros percentiles en relación con las pruebas; son valiosos como base de comparación de las calificaciones individuales con otras de una distribución.

Para la mayor parte de los fines, la media es la mejor medida de la centralización; es la única de las tres que depende del valor numérico de cada punto de una distribución. Es un indicador confiable de "centralización", porque siempre identifica el "punto de equilibrio", o "centro de gravedad" de la distribución. Como la media se presta mejor al cómputo matemático, es más conveniente para derivar otras medidas estadísticas. Por ejemplo, las medias de dos conjuntos de datos pueden usarse para computar una media para el conjunto combinado de éstos, lo que no puede hacerse con la moda o la mediana.

Empero, la media sólo puede suministrarnos información sobre el punto central de una distribución. Para entender en forma más completa un conjunto de puntos, también tendremos que saber cómo se dispersan en torno de este punto central. Por eso, los estadígrafos desarrollan medidas de *dispersión* o *variabilidad*.

La medida de distribución más simple es la extensión, que se define como la diferencia entre los puntos más altos y más bajos de una distribución: el rango de calificaciones por ortografía, de la Tabla II, se identifica fácilmente. Simplemente, restamos el más bajo (12) del más alto (20). Como la extensión sólo es sensible a los dos puntos extremos de una distribución, no se la considera como medida muy satisfactoria de la dispersión. Su debilidad está vívidamente representada en las dos distribuciones cuyos gráficos se presentan en las Figuras 4 y 5. Tanto la media (15) como la extensión (20) son idénticas para estas dos distribuciones, pero éstas son evidentemente diferentes.

Así como para la mediana y la modalidad, la debilidad de la extensión radica en que no toma en consideración el valor numérico de cada punto. Una medida natural de la dispersión,



5. En esta gráfica de distribución, la media (15) y el recorrido (20) son idénticos a los de la Figura 4; pero las distribuciones son diferentes.

que comprende a todos los puntos, es la diferencia promedio entre los individuales y su media. Como hemos visto, la media sirve como "punto de equilibrio" para una distribución; en consecuencia, podemos medir las desviaciones respecto a la media en términos de diferencias positivas y negativas. Las desviaciones en una dirección serían positivas; en la otra, negativas.

Como la media es el "punto de equilibrio", o "centro de gravedad" para una distribución, la suma de las desviaciones con respecto a ella es 0. (Esto se debe a que algunas de las diferencias son positivas, y otras, negativas.) Por esta razón, debemos usar la noción matemática de *valor absoluto* para tratar las desviaciones respecto a la media como distancias, sin considerar la dirección, esto es, sin indicar si son positivas o negativas. El promedio de los valores absolutos de las diferencias entre los puntos individuales y la media se denomina *desviación media*, y es una medida simple y precisa de la dispersión. Para la distribución representada en la Figura 4, el cálculo de la desviación media sería como sigue: $15 \cdot |5 - 15| + 10 \cdot |10 - 15| + 5 \cdot |15 - 15| + 10 \cdot |20 - 15| + 15 \cdot |25 - 15|$ dividido por 55 (frecuencia). Sería igual a $15 \cdot 10 + 10 \cdot 5 + 5 \cdot 0 + 10 \cdot 5 + 15 \cdot 10$ dividido por 55. Esto nos daría:

$$\frac{400}{55} \approx 7{,}3$$

Un cálculo similar para la distribución de la Figura 5 daría una desviación media de 4,0. Así, la cifra de la Figura 4 indica una mayor dispersión.

El uso de valores absolutos presenta dificultades matemáticas que pueden evitarse usando otra medida. Los signos positivo y negativo que llevaron a la introducción de valores absolutos podrían también eliminarse elevando al cuadrado las desviaciones con respecto a la media, puesto que el cuadrado de un número positivo, o negativo, es siempre positivo. Tal procedimiento mantiene las cualidades descriptivas de la desviación media, al tiempo que brinda una medida que es más fácil de manejar matemáticamente. De aquí que los estadígrafos prefieran usar la *desviación estándar*, que se indica por la letra griega σ, como medida de la dispersión. La desviación estándar se define como la raíz cuadrada del cuadrado de la desviación promedio respecto de la media. Entonces, para calcularla hallamos primero el promedio de los cuadrados de las desviaciones; este número se llama *variancia*. Supongamos que deseáramos encontrar la variancia y desviación estándar (σ) para los datos de la Tabla II; la obtendríamos así: $(20 - 15{,}7)^2 + 2(19 - 15{,}7)^2 + \ldots 3(13 - 15{,}7)^2 + 2(12 - 15{,}7)^2$ dividido por 35 (número total de calificaciones), lo que sería igual a 128,15 dividido por $35 \approx 3{,}66$. La variancia es, pues, 3,66. Tendríamos entonces:

$$\sigma = \sqrt{\text{variancia}} = \sqrt{3{,}66} \approx 1{,}9$$

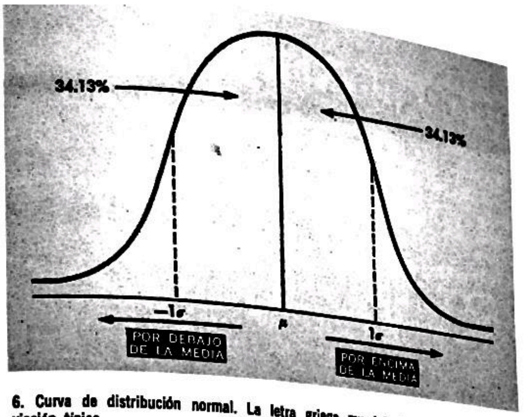
Cálculos similares darían una desviación típica de 4,2 para los datos de la Tabla III. Si comparamos estas dos medidas de la dispersión, vemos que las calificaciones de la Tabla II (muestra) no están tan "diseminadas" como las de la Tabla III (población).

INTERPRETACIÓN DE LOS DATOS

Juntas, la media y la desviación estándar nos dan una imagen razonablemente clara de una distribución, pues describen tanto su centralización como su dispersión. A veces, si conocemos la naturaleza general de la distribución, sólo necesitamos estos dos números para reconstruirla. Por ejemplo: muchos conjuntos de mediciones tienen la distribución *normal* que se muestra en la Figura 6.

Cuando un conjunto de números se ajusta a tal distribución típica, podemos determinar aproximadamente cuántos de los números caen dentro de una distancia dada de la media. Para la distribución normal, aproximadamente dos tercios de los puntos caen dentro de una desviación típica de la media; así, dada la media (15) y la desviación estándar (4,2) para el conjunto de calificaciones por ortografía de la Tabla III, podríamos predecir que aproximadamente dos tercios de éstas estarían entre 10,8 y 19,2. Como la distribución está dada de hecho en la Tabla III, vemos que 144 de las 200 calificaciones efectivamente caen en este intervalo. La predicción es bastante precisa, ya que el número de calificaciones es relativamente grande y éstas se ajustan, en efecto, a la distribución normal.

El conocimiento de tales modelos generales para distribución, unido a las leyes de la pro-



6. Curva de distribución normal. La letra griega my (μ) significa media; σ, desviación típica.

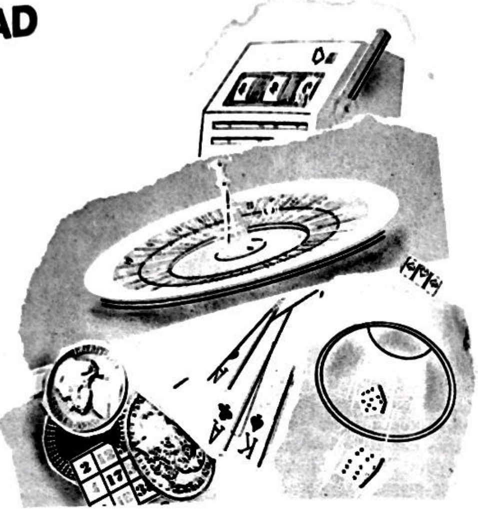
babilidad, forman la base de la *estadística predictiva*. Tanto la estadística como la probabilidad se ocupan de las distribuciones, y aprovechamos este enfoque común en la estadística predictiva. En probabilidad, se conoce el espacio de muestra (población), y se predice la composición de un conjunto de resultados (muestra). En inferencia estadística, se conoce la muestra (conjunto de resultados), e inferimos la composición de la población (espacio de muestra): así, la estadística predictiva –también llamada interferencia estadística– puede tomarse como una aplicación inversa de las leyes de la probabilidad.

Si pudiéramos estar seguros de que la distribución de puntos en una muestra refleja exactamente la distribución de los de la población de la cual se extrajo, las inferencias estadísticas serían exactas y fáciles; pero, inclusive cuando se conoce una población, la teoría de la probabilidad nos dice que las muestras no siempre serán iguales. Lo mejor que podemos esperar es que, si la muestra es suficientemente grande y escogida cuidadosamente, las características de ésta se aproximarán mucho a las de la población de donde se tomó.

Supongamos que sólo tuviésemos los datos registrados en la Tabla II y deseáramos estimar la calificación media de ortografía para todas las clases de inglés: nuestra estimación mejor sería la media obtenida para la muestra, o sea, 15,7; pero sabiendo que las muestras varían, compensaríamos esta estimación recurriendo a un *intervalo de confianza*, dentro del cual esperaríamos que cayera la media verdadera de la población. Al suponer que el conjunto total de calificaciones de ortografía está normalmente distribuido, y aplicando las leyes básicas de la probabilidad, podríamos determinar que existe una posibilidad del 95 % de que la media de población caiga en un intervalo dado con centro en 15,7. El 95 por ciento es una medida de la confianza, o credibilidad, que podemos poner en nuestra estimación de intervalos. Significa que 95 de cada 100 poblaciones de las que podría escogerse una muestra con las características dadas tendrían una media comprendida dentro del intervalo determinado. A causa de las incertidumbres presentes en el muestreo, tal estimación de intervalos, acompañada por una enunciación del grado de confianza que podemos dar a esa estimación, es preferible a una aproximación numérica simple.

El proceso de establecer intervalos de confianza nos permite probar hipótesis sobre una población. Las modificaciones de este proceso nos permiten realizar comparaciones estadísticas de dos muestras extraídas de la misma población, para comparar una muestra con una población conocida, o para inferir otras características de una población desconocida.

PROBABILIDAD



¿Cuán a menudo podemos esperar que los tres hijos de una familia sean varones?

¿Qué posibilidades tiene usted de ganar en una lotería, o de que el número que le dan al entrar a ciertas fiestas sea el premiado?

¿Cuán factible es que usted alcance el "Entablado" en un juego de Monopolio, cuando le toque su turno?

Si un bateador logra un promedio de 3 aciertos en 10 bateos, ¿cuáles son sus posibilidades de obtener 4 aciertos seguidos?

¿Se ha preguntado el lector alguna vez cuáles son las respuestas a tales preguntas? Los científicos tienen, a menudo, ocasión de estudiar preguntas de este tipo general (si bien mucho más complicadas que las cuatro anteriores) y al buscar las respuestas aplican la teoría matemática de la probabilidad.

La probabilidad es el modo que tiene el matemático para describir la factibilidad de que ocurra cierto suceso. Se la emplea para predecir el resultado de un experimento, que está gobernado por las leyes del azar. La teoría de la probabilidad nos permite determinar las características posibles de una muestra extraída de una población cuyas características son conocidas. También constituye la base de la parte de la ciencia de la estadística con ella relacionada. La inferencia estadística es una extensión de la teoría de la probabilidad en la que se usa el resultado de un experimento para estimar las condiciones que lo rigen. Estaríamos extrayendo una inferencia estadística si debiéramos hacer una conjetura condicionada sobre una población, mediante el examen de las características de una muestra que se tomó de aquélla.

Consideremos ahora algunos ejemplos. Supongamos que sabemos que hay 400 varones y 200 mujeres en cierto colegio, y llevamos a cabo el experimento de elegir al azar un estudiante de esta escuela. En una selección aleatoria, cada estudiante tendría exactamente la misma posibilidad de ser escogido. En este experimento, diríamos que la probabilidad de escoger un alumno es de $\frac{2}{3}$, y la de elegir una alumna, de $\frac{1}{3}$. Los números $\frac{2}{3}$ y $\frac{1}{3}$ indican que, en promedio, deberíamos esperar escoger un va-

rón, dos de cada tres veces; y una mujer, una de cada tres.

Si repitiésemos el experimento anterior quince veces, nuestra estimación mejor sería que la muestra elegida contendría $\frac{2}{3}$ de muchachos y $\frac{1}{3}$ de muchachas. Éste es un ejemplo de cómo se emplea la probabilidad para predecir el resultado de un experimento. La composición probable de la muestra se determinó a partir de la composición conocida de la población de la cual se seleccionó.

Supongamos que, en una segunda escuela, desconocemos la proporción de varones y mujeres; sabemos solamente que el número total de estudiantes es 600. Para estimar la relación de muchachos a muchachas podríamos realizar un experimento consistente en seleccionar 20 estudiantes al azar. Si la muestra escogida en este experimento contiene 10 alumnos y 10 alumnas, nuestra estimación más justa sería decir que en el colegio hay igual número de varones y mujeres, o sea, 300 de cada sexo. En este caso, la composición dē la muestra nos permite hacer una inferencia estadística sobre la composición de la población de donde se extrajo.

Como se ve por estos ejemplos, probabilidad e inferencia estadística están muy íntimamente relacionadas; esta última puede considerarse como una aplicación de la probabilidad en la que se invierte el proceso de razonamiento. La inferencia estadística sólo es un aspecto del campo de la estadística, que también incluye la recopilación, presentación y análisis de los datos.

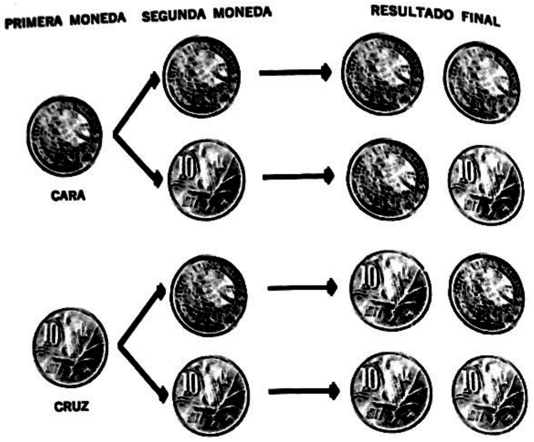
Los ejemplos anteriores ilustran la idea básica de la probabilidad y sugieren cómo puede aplicársela, pero también pueden inducir a error; se los simplificó demasiado para evitar algunos de los problemas más difíciles que surgen en el estudio de los fenómenos aleatorios. En el primer ejemplo, si repitiésemos el experimento muchas veces, podría esperarse elegir al azar un alumno en vez de una alumna unas dos veces de cada tres; pero las posibilidades de seleccionar exactamente 10 varones en una muestra dada de 15 estudiantes realmente son muy pocas; esto sólo acontecería alrededor de dos veces en diez intentos cualesquiera. La mayoría de las muestras de 15 estaría próxima a la relación teórica de $\frac{2}{3}$ (11 varones y 4 mujeres ó 9 varones y 6 mujeres, por ejemplo), pero algunas ni siquiera se aproximarían. Inclusive, podríamos escoger al azar una muestra de 15, íntegramente constituida por muchachos. En otras palabras, el resultado más probable puede no ser muy probable en modo alguno.

En el segundo ejemplo, si nuestra muestra tuviera un número igual de varones y niñas, podría llevarnos a una estimación muy deficiente de su proporción en la escuela. Más aún, podríamos haber extraído tal muestra de la primera escuela, donde había el doble de alumnos que de alumnas. En ese caso, habríamos estimado la probabilidad de elegir un muchacho como de $\frac{1}{2}$, cuando realmente era de $\frac{2}{3}$.

Muchos problemas interesantes y útiles en probabilidad son complicados, pues existen múltiples resultados posibles; si el número de éstos es infinito, debemos usar, en nuestro análisis, el cálculo infinitesimal. En esta breve introducción a la teoría de probabilidades nos ocuparemos de experimentos sencillos que producen un número de resultados relativamente reducido. Empero, los principios básicos de la probabilidad son los mismos para problemas simples que para los más complicados. A medida que usted lea, trate de entender estos principios para poder aplicarlos a situaciones más complejas; el resolver las situaciones, así como los problemas especiales, que se dan en el texto, lo ayudarán a comprender la materia y se dará cuenta de que las matemáticas no son un deporte para espectadores.

PROBABILIDAD Y JUEGOS DE AZAR

Los juegos de azar que incluyen monedas, barajas y dados nos brindan experimentos simples que producen un corto número de resultados. El uso de tales experimentos para ilustrar los principios de la probabilidad es históricamente apropiado. Los historiadores nos cuentan que un jugador del siglo XVII, el Chevalier de Méré, estaba interesado en las probabilidades implícitas en un juego de azar que se jugaba con dados. Decidió ponerse en contacto con un famoso matemático y científico, Blaise Pascal, para que éste pudiera ayudarlo con sus cálculos. Pascal quedó intrigado con los interesantes planteos que surgieron del estudio del problema de De Méré. Inició con otros matemáticos correspondencia referente a este asunto, lo que llevó al desarrollo de la teoría de la probabilidad.



1. Resultados posibles al arrojar dos monedas. En el texto aparece un análisis de la posibilidad de cada uno de los cuatro resultados probables.

Consideremos el experimento sencillo de arrojar monedas; cuando lo hacemos con una sola, hay dos resultados posibles: cara o cruz. Desecharíamos cualquier tirada en que la moneda cayese de canto. Si está en buenas condiciones y la arrojamos con vigor, parece razonable decir que cara y cruz son resultados igualmente probables. Un matemático diría que la probabilidad de cara es $\frac{1}{2}$, o sea que, en promedio, la moneda caería en cara una de cada dos veces.

Supóngase que complicamos ligeramente este experimento arrojando la moneda dos veces, o arrojando dos monedas, lo que viene a ser casi lo mismo. Vemos que podrían ocurrir tres cosas en este experimento: o bien lograríamos dos caras, a una cara y una cruz, o dos cruces. El diagrama de la Figura 1 ilustra los resultados posibles de este experimento.

Sea que CR simbolice "cara" y CZ, "cruz".

Si examinamos cuidadosamente el diagrama vemos que, realmente, hay cuatro resultados individuales –o eventos elementales– posibles: CRCR, CRCZ, CZCR y CZCZ, cuando arrojamos dos monedas. Aquí, a CRCZ no se considera igual a ZCR; podríamos, por ejemplo, usar una moneda de 5 centavos y otra de 10.

Como cada uno de los cuatro resultados individuales tiene iguales probabilidades de ocurrir, esperaríamos obtener CRCR una vez de cada cuatro; en otras palabras, la probabilidad es obtener dos caras una de cada cuatro veces o, tal como lo indicaría un matemático, $P\ (CRCR) = \frac{1}{4}$. Igualmente, diríamos $P\ (CRCZ) = \frac{1}{4}$; $P\ (CZCR) = \frac{1}{4}$; $P\ (CZCZ) = \frac{1}{4}$.

Como lo que ocurre a la primera moneda no tiene efecto alguno sobre lo que ocurre a la segunda, decimos que las dos tiradas son independientes. Cuando dos sucesos lo son, podemos emplear sus probabilidades individuales para computar la probabilidad de que ambos ocurran en un solo ensayo de un experimento. En este caso, podríamos haber usado las probabilidades que resultan de tirar una sola moneda: es decir, $P\ (CR) = \frac{1}{2}$; $P\ (CZ) = \frac{1}{2}$, a fin de computar las probabilidades que surgen al tirar dos monedas. Por ejemplo, $P\ (CRCZ) = P\ (CR) \cdot P\ (CZ) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$; $P\ (CRCR) = P\ (CR) \cdot P\ (CR) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$; el símbolo $\cdot$ significa "por". Como el suceso "una cara y una cruz" podría ocurrir de dos maneras –ya sea CRCZ o bien CZCR– podríamos esperar obtener "una cara y una cruz" dos veces de cada cuatro, en promedio, y asignaríamos probabilidad a este suceso.

También podríamos emplear las probabilidades individuales asignadas a los resultados CRCZ y CZCR, para computar la probabilidad de que ocurran ora CRCZ, ora CZCR. Cuando dos eventos no pueden ocurrir juntos en un solo ensayo de un experimento, la probabilidad de que ocurra uno u otro es la suma de sus probabilidades individuales; así, $P\ (CRCZ \text{ o } CZCR) = P\ (CRCZ) + P\ (CZCR) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. En estos ejemplos obtenemos la probabilidad de un suceso, ya sea contando los resultados de un experimento o usando probabilidades previamente determinadas.

Podemos aplicar los mismos principios en otros experimentos. Consideremos la pregunta "¿Cuán a menudo podemos esperar que los tres hijos de una familia sean varones?" Si supusiéramos que nace igual número de varones y mujeres (lo que no es totalmente cierto), diríamos que la probabilidad de un varón, $P\ (V)$ es de $\frac{1}{2}$ y la de una mujer $P\ (M)$ también es $\frac{1}{2}$. Los resultados posibles serían iguales un número, a los de arrojar tres monedas: VVV, VVM, VMV, VMM, MVV, MVM, MMV y MMM. Como VVV ocurre en sólo uno de los ocho resultados posibles, diríamos que $P\ (VVV) = \frac{1}{8}$. También podríamos haber usado la regla para computar la probabilidad de que ocurra la totalidad de varios eventos independientes: $P\ (VVV) = P\ (V) \cdot P\ (V) \cdot P\ (V) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$. De cualquier modo, llegaríamos a la conclusión de que deberíamos esperar que los tres hijos de una familia sean varones alrededor de una vez de cada ocho.

En los ejemplos tratados antes se incluyeron algunos de los principios básicos de la probabilidad. Examinemos ahora estos principios con mayor detenimiento.

ESPACIOS DE MUESTRA

Para determinar la probabilidad de un resultado particular de un experimento debemos poder identificar todos los resultados posibles: a ese conjunto de resultados posibles los matemáticos lo llaman *espacio de muestra* del experimento. En los ejemplos sencillos que pasaremos a considerar, usualmente estimaremos conveniente enumerar los espacios de muestra.

Cuando examinamos el experimento consistente en arrojar una sola moneda, nuestro espacio de muestra consistió en dos resultados posibles, CR y CZ (cara y cruz). Cuando extendimos el experimento a la tirada de dos monedas (o dos veces la misma), usamos un espacio de muestra de CRCR, CRCZ, CZCR, CZCZ. En el experimento de escoger un solo estudiante de una población escolar conocida, el espacio de muestra era el conjunto de todos los estudiantes de la escuela. Puesto que sólo nos interesaba si elegíamos a un varón o a una mujer, y no cuál de ellos, podríamos haber empleado un espacio de muestra de sólo dos elementos, Muchacho y Niña: en este caso, V pudo haberse utilizado para designar el conjunto de varones, y M, el de mujeres. Cada resultado individual de un espacio de muestra se denomina también *punto*, o *suceso elemental*.

Si bien un solo experimento producirá solamente un conjunto de resultados individuales, podremos utilizar uno cualquiera de varios espacios de muestra, según lo que estemos investigando. Por ejemplo, consideremos el experimento consistente en extraer un naipe de una baraja francesa bien mezclada; si sólo nos interesa la carta individual que se sacó, podríamos recurrir a un espacio de muestra consistente en 52 elementos: cada carta individual del mazo. En el mismo experimento, podría interesarnos sólo el valor nominal del naipe que se muestra; en este caso, nuestro espacio de muestra consistiría solamente en los 13 elementos (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, Valet, Reina, Rey), o podríamos estar interesados únicamente en el palo que se saque. En este caso, usaríamos el espacio de muestra de Tréboles, Diamantes, Corazones, Piques. Si sólo nos interesase el color, podríamos usar un espacio de muestra de sólo dos elementos (Rojo, Negro). En la mayoría de los experimentos, usamos el espacio de muestra que identifica la característica sobre la cual deseamos concentrarnos. Podemos utilizar el de los sucesos elementales para estructurar otros espacios de muestra, reuniendo los sucesos elementales con características análogas.

Probablemente ustedes hayan participado en juegos tales como Monopolio y Parcheesi, donde las jugadas estaban determinadas por la tirada de uno, o un par, de dados. Un espacio de muestra para el experimento consistente en arrojar un solo dado sería el conjunto {1, 2, 3, 4, 5, 6}. Es bastante más difícil generar un espacio de muestra para la tirada de un par de dados. Con el fin de mantener las cosas en orden, supongamos que un dado es rojo y el otro, verde; un resultado de 4 en el rojo y 3 en el verde serían diferentes al resultado de 3 en el rojo y 4 en el verde. Si acordamos anotar el resultado del dado rojo primero, y el del verde segundo, podríamos mostrar esta diferencia determinando los pares (4,3) y (3,4).

TABLA 1

A la derecha hay un espacio de muestra para la tirada de un par de dados. Se supone que uno de los dados es rojo y el otro, verde; asimismo, que el resultado del rojo se da primero en cada uno de los pares que se muestra en la Tabla. Obsérvese que, en el texto, los dos resultados –separados por una coma– se dan entre paréntesis: así, nos referiremos a (2,5), (4,3), (6,1), etcétera.

RESULTADO DEL DADO ROJO \ RESULTADO DEL DADO VERDE	1	2	3	4	5	6
1	1 / 1	1 / 2	1 / 3	1 / 4	1 / 5	1 / 6
2	2 / 1	2 / 2	2 / 3	2 / 4	2 / 5	2 / 6
3	3 / 1	3 / 2	3 / 3	3 / 4	3 / 5	3 / 6
4	4 / 1	4 / 2	4 / 3	4 / 4	4 / 5	4 / 6
5	5 / 1	5 / 2	5 / 3	5 / 4	5 / 5	5 / 6
6	6 / 1	6 / 2	6 / 3	6 / 4	6 / 5	6 / 6

La Tabla 1 muestra el espacio de muestra para el experimento de arrojar un par de dados. Supongamos que lo hacemos, y que el dado rojo cae en 4 (es decir, queda con el número 4 hacia arriba), y el verde, en 3; si, partiendo del 4 en el extremo izquierdo de la Tabla, nos corremos hacia la derecha hasta encontrar la columna encabezada por 3, hallaríamos el par (4,3). Tendremos ocasión de referirnos varias veces a este espacio de muestra en el curso de este artículo.

Veamos si es posible enumerar espacios de muestra para los experimentos siguientes:

1. Arrojar una moneda y un dado. [Sugerencia: dos puntos de este espacio de muestra podrían ser (CR,3) y (CZ,2).]

2. Arrojar una moneda de cinco, una de diez y otra de veinticinco centavos de dólar: si la de cinco sale cara; la de diez, cruz, y la de veinticinco, cara, podemos indicarlo mediante (CR, CZ, CR). Deberíamos encontrar 8 puntos para este espacio de muestra.

PROBABILIDAD DE UN SUCESO

Por definición, un *suceso* es cualquier subconjunto de un espacio de muestra; en otras palabras, si cada miembro del conjunto A lo es también del B, decimos que A es subconjunto de B. Si el conjunto A = {1, 2, 3}, y el conjunto B = {1, 2, 3, 4, 5}, entonces el conjunto A es subconjunto del conjunto B. Así, un suceso es un conjunto de resultados individuales de un experimento. Tal como hemos indicado, un suceso elemental es un solo resultado individual. Con el objeto de asignar probabilidades a un suceso –es decir, a un conjunto de resultados individuales– tendremos que poder hacerlo primero con los resultados individuales del experimento.

Por ejemplo, si consideramos el experimento de extraer una sola carta de un mazo, existen 52 sucesos elementales, o resultados individuales, posibles. Si tomamos una al azar, cada carta tiene exactamente la misma posibilidad de ser extraída, por lo que a cada uno de los 52 resultados posibles les asignaríamos la misma probabilidad: $1/52$. Obsérvese que la suma de las probabilidades asignadas al suceso elemental es 1. También lo es la probabilidad de un suceso cierto.

El mismo principio básico puede emplearse para responder a la pregunta: "¿Cuáles son sus posibilidades de ganar en una lotería, o el premio de una fiesta?" En cada caso, si el total de números es n y usted sólo tiene uno, tendrá sólo una posibilidad en n de ganar; luego, la probabilidad será 1/n.

En el experimento de extraer una carta de un mazo, quizá sólo podría interesarnos si fuese un as. Entonces, el suceso que nos interesa consiste en cuatro sucesos elementales: el as de tréboles, el de diamantes, el de corazones y el de piques. Aun así, tendríamos que extraerlos al azar de todo el mazo de cartas, que cuenta 52. Como 4 de los 52 naipes pertenecen al suceso recién descrito (4 ases), diríamos que la probabilidad de salida de un as es $4/52$. Podríamos usar los símbolos P (As) = $4/52$ para representar esta aseveración.

Supongamos que, en el experimento anterior, sólo nos interesa el color de la carta elegida: ¿Puede usted ver por qué P (Rojo) = $26/52$? Podemos emplear las probabilidades asignadas a los sucesos elementales para atribuirlas a los puntos de otros espacios de muestra. Por ejemplo: si estuviésemos usando el espacio de muestra Tréboles, Diamantes, Corazones, Piques, podríamos asignar la probabilidad de $1/4$ a cada punto de este espacio, puesto que P (Tréboles) = $13/52 = 1/4$.

Podemos resumir lo anterior, estableciendo la regla siguiente:

Si un experimento puede dar n resultados diferentes, pero igualmente factibles, y si m de ellos corresponden al suceso X, entonces la probabilidad del suceso es P (X) = m/n. (Regla 1)

Mediante la aplicación de esta regla, encontremos las probabilidades de algunos sucesos en el experimento de arrojar un par de dados. La Tabla I registra 36 resultados posibles para este experimento. Así que podríamos asignar una probabilidad de $1/36$ a cada uno de los sucesos elementales. Consideremos ahora el suceso "La suma de los números que se muestran es 7". Si r representa el número del dado rojo, y v el del dado verde, podemos representar el suceso como $r + v = 7$. ¿Cuántos de los sucesos elementales corresponden a esta cantidad?; en otras palabras, ¿cuántos de los pares de la Tabla suman 7?; al consultarla para hallar la respuesta se verá que P $(r + v = 7) = 6/36$. Si consideráramos el suceso, "El mismo número aparece en ambos dados", podríamos llamarlo "$r = v$" o, como es común en muchos juegos de dados, "doble". Existen seis de tales pares: (1,1), (2,2), (3,3), (4,4), (5,5), (6,6). Así, P $(r = v)$ = P (doble) = $6/36$. ¿Qué es P $(r + v > 7)$?; el símbolo > significa "mayor que". ¿Pueden hallar 15 puntos en el espacio de muestra que corresponde a este suceso? Tendría que poder, pues P $(r + v > 7) = 15/36$.

Estos ejemplos deberían sugerir una manera de responder a la pregunta "¿Cuán factible es que usted alcance el 'Entablado' en un juego de Monopolio, cuando le toque su turno?" Supongamos que, por ejemplo, están situados en la "Avenida Pennsylvania", que se halla 5 espacios de "Entablado"; para determinar la probabilidad de que caigan en "Entablado" realice el experimento de arrojar dos dados. Han de llegar a "Entablado" si los números que muestran los dados suman 5. ¿Cuál es P $(r + v = 5)$?

Intente resolver los siguientes problemas:

1. Si una caja contiene 4 canicas rojas y 5 blancas, ¿cuál es la probabilidad de extraer una roja al primer intento?

2. ¿Cuál es la probabilidad de que extraiga una figura (Valet, Reina, Rey) de una baraja?

3. ¿Cuál es la probabilidad de que obtenga un 5 al arrojar un solo dado?

4. ¿Cuál es la probabilidad de que no obtenga un 5 al arrojar un solo dado?

SUCESOS COMPLEMENTARIOS

Existe una relación que simplifica frecuentemente el cómputo de una probabilidad. Consideremos el ejemplo representado en la Figura 2; hay 10 puntos en el espacio de muestra E. El suceso A contiene 4 de ellos, y los 6 puntos restantes no están dentro de A. Al conjunto de puntos exteriores a un conjunto dado A se lo llama complemento de $\bar{A}$, y se lo indica en forma usual mediante el símbolo A. Si los sucesos elementales del espacio de muestra E son igualmente posibles, entonces P (A) = $4/10$ y P $(\bar{A}) = 6/10$.

En un espacio de muestra de n sucesos igualmente factibles, si uno A ocurre en m resultados, entonces el suceso $\bar{A}$ lo hará en $n - m$; así, si P $(\bar{A}) = m/n$, entonces P $(A) = (n - m)/n$. Ahora $(n - m)/n = 1 = m/n$. Obtenemos este resultado dividiendo numerador y

TABLA 1

A la derecha hay un espacio de muestra para la tirada de un par de dados. Se supone que uno de los dados es rojo y el otro, verde; asimismo, que el resultado del rojo se da primero en cada uno de los pares que se muestra en la Tabla. Obsérvese que, en el texto, los dos resultados –separados por una coma– se dan entre paréntesis: así, nos referiremos a (2,5), (4,3), (6,1), etcétera.

RESULTADO DEL DADO ROJO	RESULTADO DEL DADO VERDE 1	2	3	4	5	6
1	1 / 1	1 / 2	1 / 3	1 / 4	1 / 5	1 / 6
2	2 / 1	2 / 2	2 / 3	2 / 4	2 / 5	2 / 6
3	3 / 1	3 / 2	3 / 3	3 / 4	3 / 5	3 / 6
4	4 / 1	4 / 2	4 / 3	4 / 4	4 / 5	4 / 6
5	5 / 1	5 / 2	5 / 3	5 / 4	5 / 5	5 / 6
6	6 / 1	6 / 2	6 / 3	6 / 4	6 / 5	6 / 6

La Tabla 1 muestra el espacio de muestra para el experimento de arrojar un par de dados. Supongamos que lo hacemos, y que el dado rojo cae en 4 (es decir, queda con el número 4 hacia arriba), y el verde, en 3; si, partiendo del 4 en el extremo izquierdo de la Tabla, nos corremos hacia la derecha hasta encontrar la columna encabezada por 3, hallaríamos el par (4,3). Tendremos ocasión de referirnos varias veces a este espacio de muestra en el curso de este artículo.

Veamos si es posible enumerar espacios de muestra para los experimentos siguientes:

1. Arrojar una moneda y un dado. [Sugerencia: dos puntos de este espacio de muestra podrían ser (CR,3) y (CZ,2).]

2. Arrojar una moneda de cinco, una de diez y otra de veinticinco centavos de dólar: si la de cinco sale cara; la de diez, cruz, y la de veinticinco, cara, podemos indicarlo mediante (CR, CZ, CR). Deberíamos encontrar 8 puntos para este espacio de muestra.

PROBABILIDAD DE UN SUCESO

Por definición, un *suceso* es cualquier subconjunto de un espacio de muestra; en otras palabras, si cada miembro del conjunto A lo es también del B, decimos que A es subconjunto de B. Si el conjunto A = {1, 2, 3}, y el conjunto B = {1, 2, 3, 4, 5}, entonces el conjunto A es subconjunto del conjunto B. Así, un suceso es un conjunto de resultados individuales de un experimento. Tal como hemos indicado, un suceso elemental es un solo resultado individual. Con el objeto de asignar probabilidades a un suceso –es decir, a un conjunto de resultados individuales– tendremos que poder hacerlo primero con los resultados individuales del experimento.

Por ejemplo, si consideramos el experimento de extraer una sola carta de un mazo, existen 52 sucesos elementales, o resultados individuales, posibles. Si tomamos una al azar, cada carta tiene exactamente la misma posibilidad de ser extraída, por lo que a cada uno de los 52 resultados posibles les asignaríamos la misma probabilidad: $1/52$. Obsérvese que la suma de las probabilidades asignadas al suceso elemental es 1. También lo es la probabilidad de un suceso cierto.

El mismo principio básico puede emplearse para responder a la pregunta: "¿Cuáles son sus posibilidades de ganar en una lotería, o el premio de una fiesta?" En cada caso, si el total de números es n y usted sólo tiene uno, tendrá sólo una posibilidad en n de ganar; luego, la probabilidad será 1/n.

En el experimento de extraer una carta de un mazo, quizá sólo podría interesarnos si fuese un as. Entonces, el suceso que nos interesa consiste en cuatro sucesos elementales: el as de tréboles, el de diamantes, el de corazones y el de piques. Aun así, tendríamos que extraerlos al azar de todo el mazo de cartas, que cuenta 52. Como 4 de los 52 naipes pertenecen al suceso recién descrito (4 ases), diríamos que la probabilidad de salida de un as es $4/52$. Podríamos usar los símbolos P (As) = $4/52$ para representar esta aseveración.

Supongamos que, en el experimento anterior, sólo nos interesa el color de la carta elegida: ¿Puede usted ver por qué P (Rojo) = $26/52$? Podemos emplear las probabilidades asignadas a los sucesos elementales para atribuirlas a los puntos de otros espacios de muestra. Por ejemplo: si estuviésemos usando el espacio de muestra Tréboles, Diamantes, Corazones, Piques, podríamos asignar la probabilidad de $1/4$ a cada punto de este espacio, puesto que P (Tréboles) = $13/52 = 1/4$.

Podemos resumir lo anterior, estableciendo la regla siguiente:

Si un experimento puede dar n resultados diferentes, pero igualmente factibles, y si m de ellos corresponden al suceso X, entonces la probabilidad del suceso es P (X) = m/n. (Regla 1)

Mediante la aplicación de esta regla, encontraremos las probabilidades de algunos sucesos en el experimento de arrojar un par de dados. La Tabla I registra 36 resultados posibles para este experimento. Así que podríamos asignar una probabilidad de $1/36$ a cada uno de los sucesos elementales. Consideremos ahora el suceso "La suma de los números que se muestran es 7". Si r representa el número del dado rojo, y v el del dado verde, podemos representar el suceso como r + v = 7. ¿Cuántos de los sucesos elementales corresponden a esta cantidad?; en otras palabras, ¿cuántos de los pares de la Tabla suman 7?; al consultarla para hallar la respuesta se verá que P (r + v = 7) = $6/36$. Si consideráramos el suceso, "El mismo número aparece en ambos dados", podríamos llamarlo "r = v" o, como es común en muchos juegos de dados, "doble". Existen seis de tales pares: (1,1), (2,2), (3,3), (4,4), (5,5), (6,6). Así, P (r = v) = P (doble) = $6/36$. ¿Qué es P (r + v > 7)?; el símbolo > significa "mayor que". ¿Pueden hallar 15 puntos en el espacio de muestra que corresponde a este suceso? Tendría que poder, pues P (r + v > 7) = $15/36$.

Estos ejemplos deberían sugerir una manera de responder a la pregunta "¿Cuán factible es que usted alcance el 'Entablado' en un juego de Monopolio, cuando le toque su turno?" Supongamos que, por ejemplo, están situados en la "Avenida Pennsylvania", que se halla 5 espacios de "Entablado"; para determinar la probabilidad de que caigan en "Entablado" realice el experimento de arrojar dos dados. Han de llegar a "Entablado" si los números que muestran los dados suman 5. ¿Cuál es P (r + v = 5)?

Intente resolver los siguientes problemas:

1. Si una caja contiene 4 canicas rojas y 5 blancas, ¿cuál es la probabilidad de extraer una roja al primer intento?

2. ¿Cuál es la probabilidad de que extraiga una figura (Valet, Reina, Rey) de una baraja?

3. ¿Cuál es la probabilidad de que obtenga un 5 al arrojar un solo dado?

4. ¿Cuál es la probabilidad de que no obtenga un 5 al arrojar un solo dado?

SUCESOS COMPLEMENTARIOS

Existe una relación que simplifica frecuentemente el cómputo de una probabilidad. Consideremos el ejemplo representado en la Figura 2; hay 10 puntos en el espacio de muestra E. El suceso A contiene 4 de ellos, y los 6 puntos restantes no están dentro de A. Al conjunto de puntos exteriores a un conjunto dado A se lo llama complemento de $\bar{A}$, y se lo indica en forma usual mediante el símbolo A. Si los sucesos elementales del espacio de muestra E son igualmente posibles, entonces P (A) = $4/10$ y P ($\bar{A}$) = $6/10$.

En un espacio de muestra de n sucesos igualmente factibles, si uno A ocurre en m resultados, entonces el suceso $\bar{A}$ lo hará en n − m; así, si P ($\bar{A}$) = m/n, entonces P (A) = (n − m)/n. Ahora (n − m)/n = 1 = m/n. Obtenemos este resultado dividiendo numerador y

denominador por n, de la manera siguiente:

$$\frac{\frac{n}{n}-\frac{m}{n}}{\frac{n}{n}} = \frac{1-\frac{m}{n}}{1} = 1-\frac{m}{n}$$

Ahora podemos dar la siguiente regla general:

$P(\bar{A}) = 1 - P(A)$ (Regla 2)

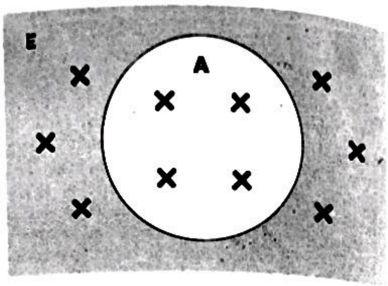
A veces es más fácil computar la probabilidad de uno de dos sucesos complementarios que la probabilidad del otro; en tales casos computamos la más fácil; para ello usamos la relación dada en la Regla 2, a fin de derivar la probabilidad del suceso complementario. Por ejemplo, en algunos juegos con dados existe un premio o un castigo cuando se echan dobles. Podríamos computar la probabilidad de no arrojar un doble, contando los puntos de muestra en la Tabla 1 que no sean dobles (hay 30); o podríamos computar la probabilidad de arrojar un doble (R = V), y emplear la relación de la Regla 2 para derivar la probabilidad de no arrojarlo (R ≠ V): el símbolo ≠ significa "no es igual".

PROBABILIDAD DE "A O B"

Una situación frecuente es hallar la probabilidad de un suceso que podría expresarse como "ya sea suceso A o suceso B". Con el suceso A o B queremos decir el conjunto de resultados que corresponde, o bien al suceso A, o al B, o, posiblemente, tanto a A como a B.

Consideremos el ejemplo de la Figura 3: existen 10 puntos en el espacio de muestra E, en el que $P(A) = 3/10$, y el de $P(B) = 2/10$. Como el suceso A o B contiene los 3 elementos de A y los 2 de B, $P(A \text{ o } B) = P(A) + P(B) = 5/10$. En este caso, no hay elementos que estén tanto en A como en B; aquí, A y B son sucesos mutuamente excluyentes, o sea, que no pueden ocurrir al mismo tiempo en un solo ensayo de un experimento. Cuando los sucesos A y B revisten tal carácter, podemos hallar la probabilidad de A o B mediante la simple suma de la probabilidad de A a la de B, tal como sigue:

Si los sucesos A y B son mutuamente excluyentes, $P(A \text{ o } B) = P(A) + P(B)$. (Regla 3)



2. Diagrama que ilustra sucesos complementarios.

Supongamos, por ejemplo, que se nos pide hallar la probabilidad de obtener ya sea un 7 ó un doble en la tirada de dos dados. Como 7 es impar, no es posible producir al mismo tiempo un 7 y un doble en una sola tirada de dos dados. Así, el suceso $(r + v = 7)$, y el suceso $(r = v)$ son mutuamente excluyentes. Ya hemos visto que $P(r + v = 7) = 6/36$, y que $P(r + v) = 6/36$: en consecuencia, aplicando la regla 3, $P(r + v = 7 \text{ ó } r = v) = 6/36 + 6/36 = 12/36$. ¿Puede verificarlo contando los puntos apropiados del espacio de muestra que se indica en la Tabla 1?

Cuando ambos sucesos no son mutuamente excluyentes –o sea que pueden ocurrir juntos en un solo resultado– no podemos usar el principio de adición de la Regla 3. Consideren la situación representada en la Figura 4. En este ejemplo, el espacio de muestra E consiste en 10 sucesos elementales igualmente factibles; 5 resultados corresponden al suceso A y 4, al B. En consecuencia, $P(A) = 5/10$, y $P(B) = 4/10$. Si aplicáramos aquí la Regla 3, tendríamos: $P(A \text{ o } B) = P(A) + P(B) = 5/10 + 4/10 = 9/10$; ésta sería una respuesta incorrecta porque los sucesos A y B no son mutuamente excluyentes, puesto que dos resultados pertenecen tanto a A como a B. Para obtener una respuesta correcta, en este caso, para $P(A \text{ o } B)$, tendríamos necesidad de aplicar la siguiente regla:

$P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B)$ (Regla 4)

Consideremos un ejemplo en el que podría usarse esta regla: supongamos que quisiésemos hallar la probabilidad de extraer una figura o un pique de un mazo de cartas. $P(\text{Pique}) = 13/52$. y $P(\text{Figura}) = 12/52$. (Recuerden que cada uno de los cuatro palos tiene tres figuras: Valet, Reina y Rey.) Hay tres naipes que son al mismo tiempo figuras y piques; luego, $P(\text{Pique y Figura}) = 3/52$; aplicando la Regla 4, podríamos arribar a la conclusión de que $P(\text{Pique o Figura}) = 13/52 + 12/52 - 3/52 = 22/52$. ¿Podría enumerar las 22 caras pertenecientes al suceso "Pique o Figura"?

Si coincidimos en que la probabilidad de un suceso que no puede ocurrir es 0, la Regla 4 podría, pues, reemplazar a la Regla 3, incluyendo los sucesos mutuamente excluyentes. He aquí por qué: cuando dos sucesos A y B son mutuamente excluyentes, entonces el suceso A y B no puede ocurrir, y diríamos que $P(A \text{ y } B) = 0$; aplicando la Regla 4, $P(A \text{ o } B) = P(A) + P(B) - P(A \text{ y } B) = P(A) + P(B) - 0 = P(A) + P(B)$, lo que sería equivalente a la Regla 3: $P(A \text{ o } B) = P(A) + P(B)$.

He aquí dos problemas bastante simples para que los intente usted:

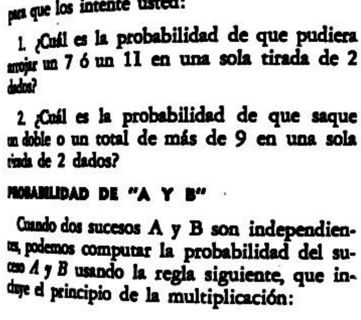
1. ¿Cuál es la probabilidad de que pudiera arrojar un 7 ó un 11 en una sola tirada de 2 dados?

2. ¿Cuál es la probabilidad de que saque un doble o un total de más de 9 en una sola tirada de 2 dados?

PROBABILIDAD DE "A Y B"

Cuando dos sucesos A y B son independientes, podemos computar la probabilidad del suceso *A y B* usando la regla siguiente, que incluye el principio de la multiplicación:

3. Diagrama de dos sucesos mutuamente excluyentes.



Si los sucesos A y B son independientes, es $P(A \text{ y } B) = P(A) \cdot P(B)$. (Regla 5)

Intuitivamente, consideraríamos a dos sucesos como independientes cuando no tienen nada que ver entre sí: si bien ésta es generalmente una regla empírica confiable, debe empleársela con cuidado. Por ejemplo, cuando dos sucesos son mutuamente excluyentes, nuestro primer pensamiento podría ser que nada tienen que ver entre sí; pero cuando uno de estos dos sucesos ocurre, no existe la posibilidad de que ocurra el otro; luego, la ocurrencia de uno de ellos ciertamente afecta la probabilidad de que también lo haga el otro. Por consiguiente, los sucesos mutuamente excluyentes nunca son independientes entre sí; para que lo sean, la ocurrencia de uno no debe afectar la probabilidad de que el otro suceda al mismo tiempo.

Aquí hay un problema que implica dos sucesos independientes, en el que podría aplicarse la Regla 5. Supongamos que arrojamos dos dados: ¿cuál es la probabilidad de que el número 1 aparezca en el dado rojo y de que el 3 lo haga en el verde? En este caso, el resultado en el dado verde nada tiene que ver con el del rojo. En consecuencia, tenemos dos sucesos independientes. La posibilidad de que salga cualquiera de los números 1, 2, 3, 4, 5 y 6 en la tirada de uno solo es 1/6. La probabilidad del resultado (1,3) al arrojar dos dados, se calcularía del modo siguiente, de acuerdo con la Regla 5: $P(1 \text{ y } 3) = P(1) \cdot P(3) = 1/6 \cdot 1/6 = 1/36$.

Si se usa el principio de la multiplicación podemos responder de varias maneras a la pregunta: "Si un bateador logra un promedio de 3 aciertos en 10 bateos, ¿cuáles son sus posibilidades de obtener 4 aciertos seguidos?" Si suponemos que los bateos son intentos independientes, podríamos computar la posibilidad de 4 aciertos seguidos de este modo: $P(AAAA) = P(A) \cdot P(A) \cdot P(A) \cdot P(A) = 3/10 \cdot 3/10 \cdot 3/10 \cdot 3/10 = 81/10.000$. Sería de esperar que el bateador lograse 4 aciertos consecutivos alrededor de 81 veces en cada 10.000 secuencias de 4 bateos. Esta probabilidad valdría si estuviésemos en lo cierto al suponer que los 4 bateos son intentos independientes, suposición que podría no ser muy sólida en este caso, pues el bateador probablemente acertaría mejor contra

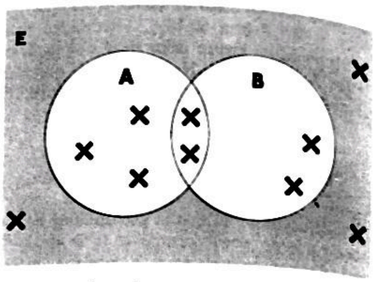
ciertos lanzadores que contra otros. Asimismo, el acertar o no también podría ejercer efecto sobre las posibilidades que tengan los próximos golpes.

Considérese el suceso "Doble y Par" en el experimento de arrojar dos dados. Por supuesto, con "Doble" queremos decir que el mismo número salga para ambos dados; por "Par", que la suma de ambos números sea un número par. En este caso, P (Doble) = $\frac{6}{36}$ y P (Par) = $\frac{18}{36}$. Si aplicamos el principio de multiplicación para sucesos independientes, tal como se dio en la Regla 5, tendríamos P (Doble y Par) = $\frac{6}{36} \cdot \frac{18}{36} = \frac{108}{1.296} = \frac{3}{36}$. Pero si examinamos la Tabla 1 encontraremos 6 puntos que corresponden al suceso Doble y Par; luego, el verdadero valor de P (Doble y Par) es $\frac{6}{36}$, y no $\frac{3}{36}$, como nuestro cálculo parecía indicar. La razón es que, en este caso, los dos sucesos no son independientes. Si tiramos un doble, estamos seguros de arrojar un número par; esto significa que, si hemos tirado un doble, la probabilidad de que también hayamos tirado un número par es 1. Por otra parte, si sabemos que hemos logrado un número par, la probabilidad de que también hayamos logrado un doble es $\frac{1}{3}$. Esto se determina contando el número de maneras en que puede salir un número par (18 en total) y cuántas de estas salidas son dobles (6).

A fin de calcular P (A y B) cuando los sucesos no son independientes, el principio de multiplicación para sucesos independientes (Regla 5) puede ser generalizado del siguiente modo:

P (A y B) = P (A) · P (B/A) (Regla 6)

B/A se lee "B dado A". Cuando sabemos que ocurre el suceso A, podemos computar P (B/A), considerando el suceso A como un espacio de muestra reducido. Para demostrar lo que quiero decir, examinemos la Figura 5: el espacio de muestra E consistente en 10 sucesos igualmente factibles, de los cuales 5 están en A, 3 en B, y 2 tanto en A como en B. Dado que ocurra A, entonces cualquier resultado obtenido es uno de los 5 de A. Si consideramos estos 5 puntos como un nuevo espacio de muestra, entonces 2 de los resultados igualmente factibles pertenecen al suceso B. Así, P (B/A) = $\frac{2}{5}$. Aplicando la Regla 6 podemos computar P (A y B) = P (A) · P (B/A) = $\frac{5}{10} \cdot \frac{2}{5} = \frac{10}{50} = \frac{2}{10}$. Como pueden ver, esto coincide con lo que obtendríamos por cuenta directa de los puntos que pertenecen tanto a A como a B en la Figura 5.



4. Diagrama de eventos; no son mutuamente excluyentes.

Si los sucesos A y B son independientes, entonces la probabilidad de que ocurra B no ha de ser afectada por el hecho de que acontece A. Así, para sucesos independientes, P (B/A) = P (B). Supongamos que queremos seleccionar una carta de un mazo: ¿cuál es la probabilidad de que la seleccionada sea al mismo tiempo una roja (R) y una figura (F)? Puesto que la mitad de la baraja es roja, sabemos que P (R) = $\frac{26}{52}$. De las 26 cartas rojas, 6 son figuras. En consecuencia, P (F/R) = $\frac{6}{26}$. Como P (F) = $\frac{12}{52} = \frac{6}{26}$, y la P (F/R) también es $\frac{6}{26}$, los sucesos F y R son independientes, y P (R y F) = P (R) · P (F). Así, la regla de multiplicación para sucesos independientes (Regla 5) es un caso especial de la Regla 6 más general: P (A y B) = P (A) · P (B/A).

He aquí algunos problemas que incluyen la probabilidad de "A y B".

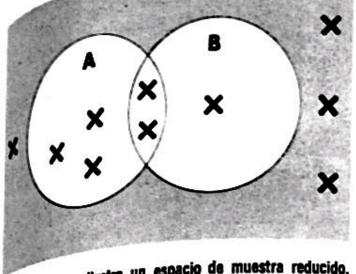
1. Si se arroja una moneda y un dado: ¿cuál es la probabilidad de obtener cara y un 3?

2. Si se extrae una carta y se arroja un dado: ¿cuál es la probabilidad de obtener dos 6?

3. De un mazo bien barajado se extrae una carta: ¿cuál es la probabilidad de que no sea ni una figura ni una roja?

GRADO DE CONFIANZA

En esta breve introducción a la probabilidad sólo hemos tocado unos pocos de sus principios básicos; el alcance de la teoría de la probabilidad es mucho más amplio que lo sugerido en nuestra exposición. Por ejemplo, simplificamos mucho las cosas al proponer únicamente situaciones donde los resultados individuales de un experimento eran igualmente factibles; en muchos casos no es así; entonces, los cálculos se hacen más complejos, y se agregan nuevas dimensiones al estudio de la probabilidad.



5. Diagrama que ilustra un espacio de muestra reducido.

En nuestra representación simplificada nos limitamos a señalar que las predicciones basadas en la probabilidad son inciertas. El científico lo acepta, pero quiere saber hasta dónde lo son; reconoce que, al tratar de predecir el resultado de un experimento sujeto a leyes del azar, a menudo puede estar equivocado. Necesita saber hasta qué punto y con qué frecuencia se equivoca; quiere establecer el grado de confianza de sus predicciones. La teoría de la probabilidad le brinda la base para determinar este grado particular de confianza.

MUCHOS USOS

Al profundizar en el estudio de las matemáticas se tendrá ocasión de trabajar con algunos de los aspectos más complicados y seductores de la teoría de la probabilidad, y se logrará una idea más clara de las muchas formas en que esta teoría sirve al hombre. Señalaré unas pocas: permite a los científicos fijar el límite dentro del cual deben mantenerse las desviaciones respecto a una ley física dada para que no contradigan dicha ley. Se la ha empleado para calcular las posiciones y velocidades de electrones que orbitan en torno de los núcleos atómicos. Mediante su aplicación se han analizado las fluctuaciones en densidad de un volumen dado de gas. Ha desempeñado un papel importante en genética; entre otras cosas, hizo posible calcular el porcentaje de individuos con rasgos similares y disímiles en generaciones sucesivas. Los fabricantes usan la teoría de la probabilidad para predecir la calidad de los artículos que salen de las líneas de producción en masa. Los expertos en seguros la usan extensamente para calcular, por ejemplo, las esperanzas de vida, de modo que puedan fijar primas de seguro de vida adecuadas. Finalmente, mediante su empleo se pueden programar computadoras electrónicas y predecir el resultado de elecciones sobre la base de comparativamente pocas urnas o máquinas escrutadas, en general con un grado de precisión sorprendente. Los fracasos comparativamente escasos de la predicción eleccionaria computada no pueden oscurecer el éxito logrado por esta técnica.

Blaise Pascal, que desarrolló la teoría de la probabilidad.

Archivo Bettmann

TEORÍA DEL JUEGO

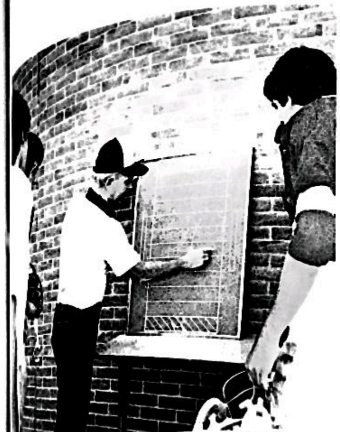
Cinco delincuentes se reúnen en un edificio, tramando un delito. Afuera aguarda un solitario policía, decidido a capturar al jefe. Los delincuentes no saben que está. No obstante, acuerdan ir saliendo sin orden ni concierto, de uno en uno, para no llamar la atención. Todo lo que el policía sabe es que el jefe es el más alto de los delincuentes. No puede capturar más que a uno de los hombres porque sería derrotado. No tiene modo de conseguir auxilio. Conforme los hombres van saliendo ¿a cuál arrestará el policía? ¿Al primero? ¿Al segundo? ¿Al último?

Si el policía hiciera el arresto al azar, sólo tendría un 20 por ciento de probabilidades de acertar con su hombre. Los teóricos del juego dicen que sería mejor que siguiera una estrategia o plan de acción: dejar escapar a los dos primeros hombres y detener al siguiente en salir que sea más alto que los otros dos. Haciéndolo así, el policía tiene un 40 por ciento de probabilidades de capturar al jefe.

Éste es un problema sencillo en la teoría del juego, un fascinante ramo nuevo de las matemáticas que trata del riesgo en situaciones conflictivas. Toma su nombre del hecho de que muchas situaciones reales difíciles en los negocios, las finanzas, la guerra, etc., son semejantes en su estructura básica a los juegos de salón que tan bien conocemos: bridge, póquer, damas, ajedrez e incluso el juego de tres en raya, gato o ta-te-tí. Tanto en estos juegos como en los acontecimientos diarios, las personas compiten unas con otras. Existen reglas según las cuales debe practicarse el juego: hay resultados, o finales –tales como ganancia, pérdida y empate– que surgen de las diferentes jugadas o estrategias de los adversarios; en general, uno o más de los jugadores interesados disponen de alguna información.

Dos niños jugando al ajedrez, un juego suma cero, que comprende un par de jugadores que siguen reglas determinadas, con un objetivo claro: simplemente ganar. Las piezas han sido hechas con objetos de quincalla.

Mimi Forsyth, Monkmeyer



Jan Braunle

Mickey Palmer, DPI

El complicado juego de fútbol norteamericano comprende dos equipos que se enfrentan para ganar un partido, con un tanteador numérico superior. Existen reglas, pero puede emplearse todo tipo de estrategias y tácticas. Izquierda: el entrenador explica una jugada a miembros del equipo. Derecha: una jugada en acción.

En la teoría del juego se supone que el objetivo de cada jugador es el de maximizar sus ganancias o minimizar sus pérdidas. También se supone que se enfrenta con un adversario sensato cuyo objetivo es hacer lo mismo. Por supuesto, la situación en "policías y ladrones" es realmente la de un juego unipersonal. Los ladrones no saben que están compitiendo con el policía. Si lo supieran, ensayarían su propia estrategia. Podrían presumir que el policía es un teórico del juego y en consecuencia hacer salir primero a su jefe. Por otra parte, el policía podría darse cuenta de que estaban pensando eso y en consecuencia detener al primer hombre que saliera. En suma, un juego puede resultar complicado.

Aparte de los juegos unipersonales, que suelen tener poco valor y ser fáciles de resolver, quizás los juegos más sencillos son los de dos personas (individuos o equipos), donde las pérdidas de uno son las ganancias del otro. Esta clase de juego se llama juego de dos personas a suma cero. Suma cero significa que lo que gana un lado el otro lo pierde, y viceversa. El valor total del juego no aumenta ni disminuye durante su curso.

Es fácil imaginar juegos donde esto no ocurra. Dos anunciantes que compiten por un mercado, por ejemplo, podrían no sólo aumentar o disminuir su participación en éste por medio de su propaganda, sino que en realidad podrían incrementar el mercado total. Para mejor ilustrar la teoría del juego nos detendremos en el juego de dos personas a suma cero.

CASO DE UN COBRO COMPLICADO

La mayoría de los juegos es finita. Es decir, que existe sólo cierto número de estrategias u opciones que cada jugador puede se-

FIGURA 1

Matriz de juego para usted y Oscar

		Estrategias de Oscar		
		A	B	
Sus estrategias	A	30	60	(30) Su maximin
	B	40	20	20
Minimax de Oscar		(40)	60	

guir, y el juego se "resuelve" en determinado número de jugadas o movimientos. Algunos juegos son también de información perfecta: sabemos o podemos ver cada movimiento que hace nuestro adversario. El ajedrez, las damas y el de tres en raya son ejemplos de juegos finitos o de información perfecta. El ajedrez es muy interesante por los muchos movimientos diferentes que admite. Un juego como el de tres en raya despierta menos interés porque tiene muy pocos movimientos y puede resolverse muy fácilmente. De hecho, en este juego el primer jugador que sale puede asegurarse por lo menos una jugada si hace su marca en uno de los cuadros de la esquina. El segundo jugador puede garantizar una jugada si hace su marca en el cuadro del centro, suponiendo, naturalmente, que cada jugada subsiguiente esté destinada a bloquear a su adversario.

Para ilustrar la teoría del juego consideremos uno finito de dos personas a suma cero, de información imperfecta. Supongamos que Juan le debe a usted \$ 60 y José le debe \$ 40. Para cobrar el dinero, usted tiene que estar en el lugar de empleo de Juan o de José el día de pago a las 17 hs. La situación se complica por el hecho de que esos dos hombres deben las mismas sumas de dinero a un individuo que llamaremos su oponente: Oscar.

Quienquiera que esté en el lugar de empleo de Juan o de José a las 17 hs. obtiene el dinero. Si los dos acuden al mismo lugar, se lo reparten. Pero si ninguno de los dos aparece, Oscar obtendrá el dinero, porque es un amigo especial de Juan y de José. Tanto Oscar como usted poseen esta información. De modo que deben determinar a quién tienen que dirigirse para conseguir la ganancia: ¿a Juan o a José?

Éste es un problema que puede resolverse utilizando la teoría del juego. Es una situación conflictiva: usted quiere obtener todo el dinero que pueda, mientras Oscar, que podría obtener todo el dinero si usted no fuera a ninguno de los dos sitios, quiere minimizar sus pérdidas. Es un juego a suma cero, porque la cantidad total que se ha de pagar (\$ 100) es la misma, cualquiera sea la estrategia empleada.

ELECCIÓN DE ESTRATEGIAS

El juego que usted y Oscar juegan se llama un juego 2 × 2: usted dispone de dos estrategias, y Oscar de otras dos. Cada uno de ustedes puede ir a ver, ya a Juan (estrategia A), ya a José (estrategia B). Algunos juegos de dos personas disponen de muchas más estrategias para los jugadores. Por ejemplo, en un juego de 2 × 3, su oponente tiene tres estrategias para elegir, mientras usted no tiene más que dos. En un juego de 4 × 3, usted tiene cuatro estrategias, contra tres de su adversario. En un juego de 3 × m, m significa que su contrario puede elegir entre muchas estrategias.

Paul Conklin. Monkmeyer

Una compradora compara marcas de productos en una tienda de comestibles. Fabricantes y anunciantes aplican la teoría del juego para captar una parte del mercado de la alimentación.

Los juegos grandes pueden reducirse a proporciones menores, porque algunas de las estrategias carecerán de sentido respecto del juego. Por ejemplo, si todos los resultados de una estrategia particular son cero, no hay razón para elegir esa estrategia; puede muy bien eliminarse.

La mejor manera de ilustrar un juego es mediante una *matriz de resultado*. Ésta indica lo que ocurre cuando los jugadores eligen sus diferentes estrategias. El cuadro 1 muestra una matriz de resultado para el juego entre usted y Oscar.

Es habitual que la matriz muestre los resultados obtenidos por el jugador cuyas estrategias están anotadas a la izquierda de la misma, quien desea maximizar sus ganancias. Los resultados son pérdidas para la persona nombrada en la parte superior de la matriz.

Ésta indica que si tanto usted como Oscar eligen la estrategia A (ir a ver a Juan), el resultado será de \$ 30. Si usted elige A y Oscar elige B (ir a ver a José), el resultado será \$ 60. Si usted elige B y Oscar elige A, el resultado será \$ 40. Si ambos deciden elegir la estrategia B, el resultado será \$ 20.

Fuera de los cuadros de la ilustración 1 se indican algunas cifras adicionales. Esto nos ayuda a "resolver" el problema siguiendo el procedimiento de la teoría del juego. La teoría particular que vamos a seguir fue primitivamente propuesta por el matemático John von Neumann en 1929 y extensamente desarrollada en un libro que escribió con Oskar Morgenstern, publicado en 1944, titulado *Teoría de los juegos y comportamiento económico*.

Las cifras que aparecen fuera de la matriz se llaman *cifras marginales*. La columna de la derecha da los mínimos de hilera: la cantidad mínima en la hilera opuesta a la que aparece. La hilera al pie de la matriz da los máximos de columna: la cantidad máxima en la columna bajo la cual aparece. La teoría del juego dice que usted debería elegir la estrategia que proporciona el máximo de sus ganancias mínimas. A esto se lo llama su *maximin*. Si su adversario es sensato, debería elegir la estrategia que proporciona el mínimo de sus pérdidas máximas. A esto se lo llama su *minimax*. En la Figura 1, su maximin y el minimax de Oscar aparecen dentro de un círculo.

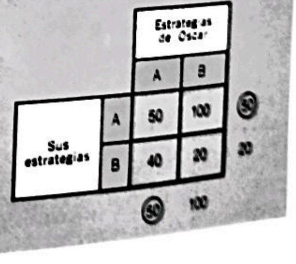
El propósito de la teoría del juego es eliminar de su problema la casualidad o la probabilidad. Escogiendo la estrategia A, usted se garantiza una ganancia de por lo menos \$ 30. Si su adversario es suficientemente incauto para elegir la estrategia B, usted ganará \$ 60. Oscar, por su parte, elegirá la estrategia A porque le garantiza una pérdida máxima de \$ 40, si usted trata de ser astuto y elige la estrategia B. Pero Oscar sabe que no va a intentar la estrategia B porque sólo obtendría \$ 20 si él también escoge B. Naturalmente, uno de los dos podría tratar de burlar al otro. Pero entonces serían unos tahúres, y esto es lo que evita la teoría del juego, que es, en cambio, una técnica matemática para garantizar una ganancia mínima (o una pérdida máxima) en una situación conflictiva.

ESTRATEGIAS MIXTAS Y PUNTOS DE EQUILIBRIO

Hemos supuesto que su juego con Oscar se jugaría sólo una vez. Es decir, que cada jugador haría un solo movimiento. Pero la mayoría de los juegos, tanto en la vida real como en los juegos de salón, comprende una cantidad de movimientos. Cara o cruz es un buen ejemplo. En este juego, si usted elige una estrategia "pura", tal como jugar a cara todo el tiempo, puede acabar perdiendo su camisa (o por lo menos todos sus centavos). Frecuentemente, tal situación reclama una estrategia "mixta", como jugar a cara y cruz cada 50 por ciento del tiempo, a la aventura. Agitando la moneda antes de cada jugada se alcanzará ese fin. En realidad, incluso aquellos que nada saben acerca de la teoría del juego, habitualmente juegan de esta manera.

FIGURA 2

Matriz de juego con punto de equilibrio



Ahora volvamos a mirar su juego con Oscar y supongamos que se ha de jugar muchas veces. ¿Tiene sentido entonces una estrategia pura? No del todo. Hemos determinado que Oscar elegirá la estrategia A porque si no lo hace usted puede ganar $ 60 en vez de $ 30. Pero vamos a engañarlo. Elijamos la estrategia B, ganando $ 40 en vez de $ 30. ¿Cuánto tiempo permitirá Oscar esto? Si usted elige B, entonces él también elegirá B, reduciendo así sus ganancias a $ 20. Puede usted volver a la estrategia A y ganar $ 60. Pero él también retornará a la estrategia A. Antes de mucho tiempo, ambos estarán empleando estrategias mixtas, tratando de ganar al otro. Ahora el problema está en saber qué clase de estrategia mixta van ustedes a emplear.

Antes de contestar a esta pregunta consideremos si debe usarse o no una estrategia mixta en un determinado juego. La teoría del juego dice que no debe utilizarse una estrategia mixta si el juego tiene un *punto de equilibrio*. Se dice que un juego tiene un punto de equilibrio si el maximin iguala al minimax; es decir, si el máximo de sus ganancias mínimas es igual al mínimo de las pérdidas máximas de su adversario. El juego con Oscar no tiene un punto de equilibrio, como puede verse comparando las dos cifras encerradas en un círculo del cuadro 1. Por consiguiente, el juego reclama una estrategia mixta. Si las dos cifras en círculo fueran iguales, entonces el juego tendría un punto de equilibrio, y cada jugador escogería la estrategia pura dictada por su maximin o minimax. Ilustremos esto.

En el cuadro 2 hemos cambiado los resultados finales con objeto de crear un punto de equilibrio para el juego. Su maximin iguala al minimax de Oscar. En estas condiciones, no tiene sentido que usted elija la estrategia B. Unicamente perdería. Oscar, quien supone que usted es sensato, no puede hacer otra cosa que elegir también la estrategia A. Si elige la estrategia B, soporta la pérdida de $ 100 en vez de $ 50. De modo que los dos seguirán eligiendo la estrategia A. Al producirse esta situación, se dice que tal juego está estrictamente determinado.

Volvamos ahora al juego original y veamos qué estrategia mixta emplearía cada jugador para maximizar sus ganancias (suponiendo que el juego se repita muchas veces).

FIGURA 3
Matriz de juego con sobrantes

		Estrategias de Oscar		
		A	B	
Sus estrategias	A	30	60	30
	B	40	20	20
		10	40	

VALOR DEL JUEGO

Todo juego tiene cierto valor. El valor del juego estrictamente determinado del cuadro 2 es de $ 50. El valor del juego original (en el cuadro 1) jugado una sola vez es $ 30. (Estos dos valores son, por supuesto, los de usted, y continuaremos hablando del valor del juego desde su punto de vista.)

Si el juego original se ha de jugar en una base continuada, tanto Oscar como usted comenzarán mezclando sus estrategias, tratando cada uno de aventajar al otro. Ambos están realmente luchando por los $ 10 repartidos entre su ganancia mínima de $ 30 y su pérdida máxima garantizada de $ 40, si cada uno adopta una estrategia única. Mezclando estrategias, cada uno espera obtener lo más posible de estos $ 10. Esto sugiere que el valor del juego está entre $ 30 y $ 40. ¿Cómo puede asegurarse Oscar de que usted no obtenga más de esa cantidad? La respuesta consiste en mezclar estrategias en una forma prescrita.

Una vez averiguado que el juego no tiene un punto de equilibrio, tache las cifras del borde. Determine las diferencias absolutas entre los resultados finales de cada hilera y columna y escriba estas cifras en los márgenes. A estas cifras se las llama *sobrantes*, o parte de los remanentes de usar cada una de las estrategias. Hemos hecho esto en el cuadro 3. Ahí puede ver que las diferencias de los resultados en la línea son de $ 30 y $ 20, mientras las diferencias de los resultados en la columna son de $ 10 y $ 40. Estas cifras son las proporciones con que deberían jugarse las diferentes estrategias (3 a 2 para usted y 4 a 1 para Oscar). Cada cifra, sin embargo, se aplica a la estrategia contraria. En otras palabras, su estrategia mixta sería emplear tres veces la estrategia B y dos la estrategia A en cada cinco jugadas; la estrategia mixta de Oscar sería recurrir a la estrategia A cuatro veces, y una a la B en igual número de jugadas.

Aunque su oponente puede fácilmente calcular la estrategia total de usted, es mejor mantenerlo tratando de adivinar cada jugada individual. De otro modo, podría embaucarlo. Usted debe elegir cada movimiento al azar. Por ejemplo, eche dentro de un sombrero tres naipes franceses rojos que representen la estrategia B, y dos naipes negros que representen la estrategia A. Antes de cada jugada mezcle las cartas y saque una. Deje que ésta determine la estrategia que elegirá. Vuelva a meterla en el sombrero y repita el proceso para la jugada siguiente. A la larga, usted empleará sus estrategias en una proporción de 3 a 2, y es grande la probabilidad de que alcance el *valor* del juego. Este valor se determina fácilmente. Para cada hilera multiplique los sobrantes por cada resultado de la otra hilera. Añada los productos y divida por la suma de los sobrantes. Halle el promedio de sus respuestas para las hileras. Esto da el valor del juego si usted emplea su prescrita estrategia mixta. Haga lo mismo con las columnas. Esto da el valor del juego si su oponente sigue su estrategia mixta prescrita. Estos dos valores deberían ser los mismos. A la larga, el valor del juego será alcanzado si ustedes dos se atienen a sus estrategias mixtas prescritas. Si uno de los jugadores se desvía de su estrategia prescrita, lo probable será que no alcance el valor del juego.

La aritmética implicada en determinar el valor del juego está ilustrada en estas ecuaciones:

$$\frac{20 \times 30 + 20 \times 60}{20 + 30} = 36$$

$$\frac{30 \times 40 + 30 \times 20}{20 + 30} = 36$$

$$\text{Valor del juego} = \frac{36 + 36}{2} = 36$$

$$\frac{40 \times 30 + 40 \times 40}{40 + 10} = 56$$

$$\frac{10 \times 60 + 10 \times 20}{40 + 10} = 16$$

$$\text{Valor del juego} = \frac{56 + 16}{2} = 36$$

El valor del juego, si ambos jugadores juegan sensatamente (o sea si usan sus prescritas estrategias mixtas) es $ 36.

…Matriz de juego para el dilema de los prisioneros

		Prisionero B	
		C	NC
Prisionero A	C	5,5	0,20
	NC	20,0	1,1

JUEGOS SUMA NO CERO Y PERSONA N

Cuando nos apartamos del juego de dos personas suma cero, los problemas se hacen un poco más complicados, porque la psicología, la negociación y la comunicación pueden convertirse en factores que los alteran.

Se recordará que un juego suma no cero es aquel en el cual las pérdidas de un jugador no son necesariamente las ganancias del otro. O sea, que el valor total del juego no se mantiene indefectiblemente siempre el mismo a lo largo de la partida. Un ejemplo de juego de dos personas suma no cero es el "Dilema de los prisioneros".

Imaginemos que dos delincuentes sospechosos del robo de un banco son detenidos y encerrados en celdas separadas para que no se puedan comunicar. Si uno de ellos confiesa su delito y declara contra su cómplice, saldrá libre. Al otro se lo condenará a 20 años de cárcel. Si ambos delincuentes confiesan y se entregan a la clemencia del tribunal, cada uno recibirá una sentencia de 5 años. Si ninguno de los dos confiesa, obtendrán 1 año cada uno por tenencia ilícita de armas. Si usted fuera uno de los prisioneros ¿qué haría?

Esta situación conflictiva se describe en el cuadro 4. Las estrategias son confesar (C) y no confesar (NC). Porque éste es un juego suma no cero, los resultados de ambos jugadores están anotados en cada cuadro, con el resultado del jugador A a la izquierda de la coma y el resultado del jugador B a la derecha. Estudie sólo su propio resultado. Como buen teórico del juego, deberá escoger la estrategia que proporciona su minimax, el mínimo de sus

pérdidas máximas. Así, deberá confesar. Lo peor que podría ocurrir sería una sentencia de 5 años.

Su adversario, el preso B, deberá elegir la misma estrategia. Pero esto podría simplificar el juego. Si piensa que su opositor es alguien que no confesará nunca, ¿qué hará, entonces? Depende de qué tipo de persona sea usted.

Supongamos que los dos son compañeros de celda y pueden discutir la situación. ¿Coincidirán ambos en no confesar? Esto tendría sentido. Pero ¿qué decir del peligro de una traición?

Como puede verse, para muchos (si no la mayoría) de los juegos, las matemáticas no tienen todas las respuestas.

Los juegos que incluyen a más de dos personas (juegos persona n) pueden ser incluso más complejos. Imaginemos una situación que incluya a tres personas. A, B y C. Si A coopera con B, se repartirán $ 6. Si A coopera con C, se repartirán $ 8. Si B coopera con C, se repartirán $ 10. Esto va ilustrado en el triángulo del cuadro 5. ¿Quién cooperará con quién para dividir el botín?

–Mire –dice A a B–, coopere conmigo y le daré $ 4, quedándome sólo con $ 2.

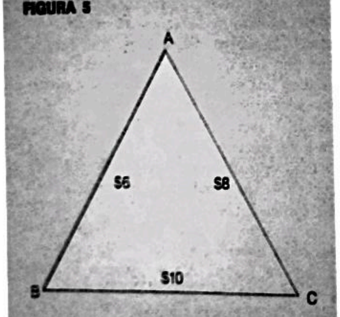
–No sea tonto –dice B a A–, probablemente conseguiré $ 5 si coopero con C.

–Nada que hacer –dice C–. Creo que puedo efectuar un trato mejor con A. ¿Qué le parece, A? ¿Tomará $ 2 y me dejará $ 6?

–Muy bien –dice A.

B se espanta.

FIGURA 5



–Mire, A –dice–. Estoy dispuesto a repartir los $ 6, mitad por mitad.

–No pierda la cabeza –dice C a B–. Le daré $ 4.

–Pero eso es lo que *yo* le ofrecí –dice A.

Siguen dando vueltas al asunto. La división del dinero dependerá del talento para negociar de los interesados. Todo lo que la teoría del juego puede decirnos es que, cooperando con uno de sus adversarios, lo más que A puede esperar recibir es $ 2; lo más que B puede esperar obtener es $ 4; y lo más que C puede esperar recibir es $ 6.

Pero cambiemos las reglas. Supongamos que A, B y C pudieran ganar un total de $ 18 si los tres cooperasen. ¿Cómo se repartiría el dinero? La resolución arriba indicada sugiere que A recibiría $\frac{1}{6}$ ó $ 3; B recibiría $\frac{1}{3}$ ó $ 6; y C recibiría $\frac{1}{2}$ ó $ 9.

Como puede verse, estos sencillos juegos dan sólo una indicación de la complejidad de más amplios juegos suma no cero. La teoría del juego ha tenido su mayor éxito hasta ahora únicamente con los juegos o situaciones conflictivas más elementales. Indudablemente, llegará a tener cada vez más importancia en el futuro, ya que gran parte de la vida tiene que ver con el conflicto y la competición. Como John D. Williams ha escrito en su libro *El estratega completo:* "El concepto de una estrategia, las distinciones entre jugadores, el papel de los acontecimientos casuales, la noción de las representaciones de los resultados en forma de matrices, los conceptos de estrategias puras y mixtas, etc., proporcionan una valiosa orientación a personas que han de reflexionar sobre situaciones conflictivas complicadas."

CÁLCULO

Una de las mayores contribuciones a las matemáticas modernas, la ciencia y la ingeniería fue la invención del *cálculo*, o, como se denomina a veces, del ***análisis matemático***, realizada aproximadamente a fines del siglo XVII. Resulta obvio decir que sin este ramo fundamental de las matemáticas muchos logros tecnológicos, tales como la llegada del hombre a la Luna, habrían sido muy difíciles, o imposibles, de alcanzar.

La palabra "cálculo" viene de la palabra latina que significa guijarro. Este nombre posiblemente se originó porque los guijarros se usaban hace miles de años para contar y resolver problemas de aritmética. Palabras similares, que a menudo utilizamos, son ***calcular***, ***calculador*** y ***calculista***.

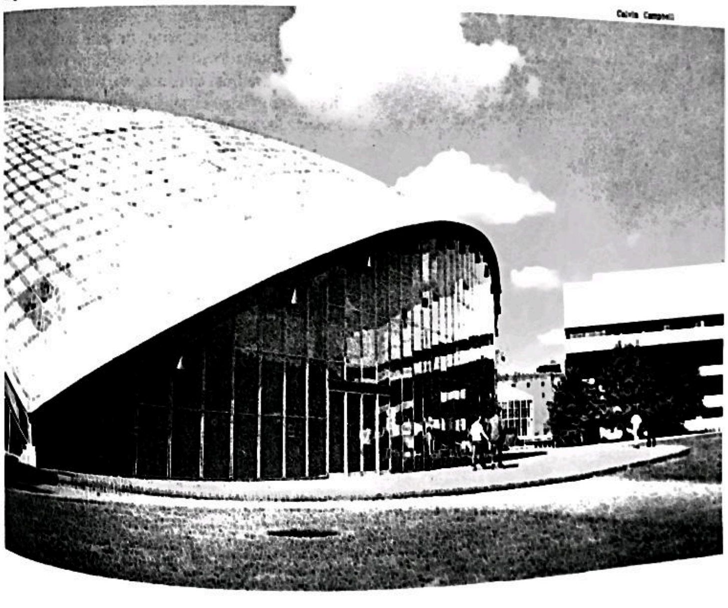
La invención del cálculo se atribuyó a dos personas que vivieron durante el siglo XVII: sir Isaac Newton, de Inglaterra, y el barón Gottfried Wilhelm von Leibniz, de Alemania. Ellos desarrollaron, independientemente, las ideas básicas del análisis matemático, con diferencia de pocos años uno del otro.

Newton, que fue uno de los más grandes físicos de todos los tiempos, aplicó el análisis matemático a sus teorías del movimiento y la gravitación. Estas teorías, frecuentemente designadas como *leyes de Newton*, le permitieron enunciar matemáticamente el movimiento de todos los objetos en el Universo, desde el lanzamiento de una pelota al aire hasta la revolución de la Tierra y los demás planetas del Sistema Solar, alrededor del Sol.

Antes de Newton y Leibniz, para resolver problemas se utilizaban las matemáticas comúnmente enseñadas en las escuelas secundarias modernas. Estas comprendían materias tales co-

Ventanas aproximadamente rectangulares limitan el área bajo el techo del Auditorio Kresge del Instituto de Tecnología de Massachusetts. En análisis matemático utilizamos rectángulos para hallar el área bajo una curva.

Calvin Campbell



mo aritmética, álgebra, geometría y trigonometría. Los principios básicos de estas disciplinas se conocieron por lo menos 1.500 años antes de Newton y Leibniz. Pese a que los principios matemáticos estudiados en estas disciplinas resultaron útiles en la resolución de cierto tipo de problemas, no servían en absoluto para resolver los relacionados con *cantidades cambiantes*, o *variables*. El análisis matemático se inventó para trabajar con cantidades cambiantes, o variables. Por tanto, podemos decir que el análisis matemático es la *matemática del cambio*.

CANTIDADES VARIABLES

Si un hombre viaja en automóvil a una velocidad de 50 kilómetros por hora, sabemos que en 2 horas recorrerá una distancia de 100 kilómetros. En la práctica, como es natural, un conductor difícilmente viaje a la misma velocidad de 50 kilómetros por hora durante 2 horas. Algunas veces se detendrá, otras irá a 80 kilómetros por hora, y a veces, a 40.

Los científicos usan el análisis para calcular la "velocidad de escape" y la velocidad y posición del cohete en un instante cualquiera.

NASA

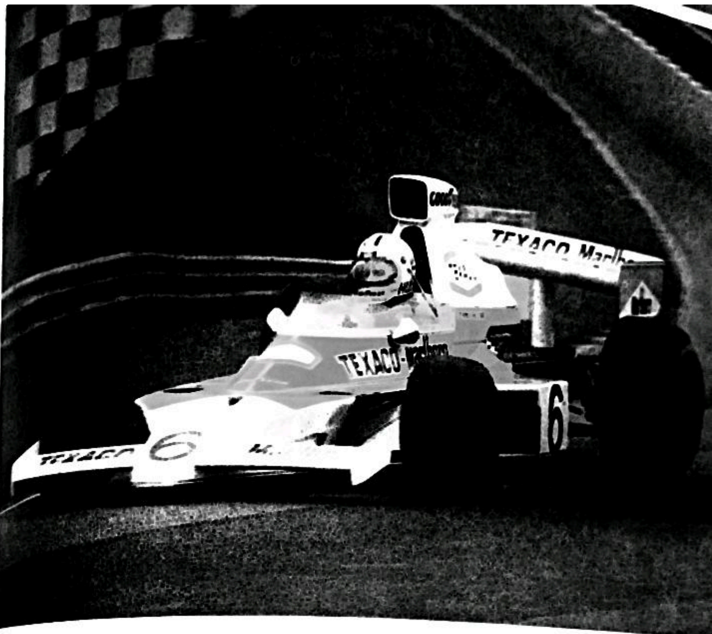
La velocidad de un automóvil es ordinariamente una cantidad cambiante, o variable. Cuando la velocidad aumenta, como al pasar de 30 a 40 kilómetros por hora, a menudo decimos que el automóvil está *acelerando* o que está realizando una *aceleración*. Del mismo modo, cuando la velocidad disminuye, o cuando el auto pasa de 40 kilómetros por hora a 20, a menudo decimos que el automóvil está *desacelerando*, o realizando una *desaceleración*. Si el automóvil mantiene la misma velocidad, decimos que está viajando a una *velocidad constante*, o *uniforme*.

Otro ejemplo de una cantidad cambiante es una pelota que se ha dejado caer, o que ha sido arrojada. Supongamos que la dejamos caer desde lo alto de un edificio. En el instante de soltarla, la velocidad es cero. Gradualmente, esa velocidad aumenta; o sea que la pelota se acelera; finalmente, cuando da contra el suelo, afirmamos que está viajando a su velocidad máxima.

Del mismo modo, si arrojamos una pelota al aire, al principio va rápido, o sea, con gran velocidad. Gradualmente aminora la marcha, o desacelera, hasta que su velocidad es cero y se detiene por un instante. En este punto la pelota ha alcanzado su altura máxima y comienza a caer. A medida que cae, su velocidad aumenta hasta que llega al suelo. El cambio en la velocidad de la pelota –o cualquier objeto que se arroja– se debe a la atracción de la Tierra: la fuerza de gravedad, o la fuerza gravitacional.

Lo mismo es válido para un cohete lanzado desde la superficie terrestre. Utilizando el análisis matemático es posible hallar la velocidad que el cohete debe tener para escapar de la gravedad terrestre, o sea, no volver a la Tierra. Esta velocidad, a menudo denominada *velocidad de escape*, resulta ser de alrededor de 11 kilómetros por segundo, o de 40.000 kilómetros por hora. Por medio del análisis matemático podemos calcular el tiempo que tardaría el cohete para llegar a la Luna y cuánto combustible se necesitaría.

Hay muchos otros ejemplos de cantidades cambiantes, o variables. Cuando una gota de lluvia o un copo de nieve caen, su tamaño aumenta gradualmente. La población de un país cambia cada año, o aun cada día. El costo de la vida cambia y también la cantidad de una sustancia radiactiva, como el radio.



La velocidad de este auto casi nunca es constante. El auto acelera al comienzo de la carrera, aminora su marcha antes de cada curva y acelera nuevamente en las rectas.

¿QUÉ ES UNA VARIABLE?

Toda cantidad que está cambiando, o variando, se denomina a menudo en matemáticas una *variable*. En los ejemplos del automóvil, la pelota y el cohete, la velocidad es una variable. También lo es la distancia de la pelota o del cohete a la superficie terrestre. Inclusive el tiempo, que está cambiando, es una variable. La población de un país o el costo de la vida son una variable. Cuando ponemos aire en nuestros neumáticos, la presión del aire cambia y es una variable. La temperatura exterior y la humedad son variables.

En matemáticas, representamos a menudo variables mediante símbolos, tales como letras del alfabeto. Así podemos hacer que v represente la velocidad, t, tiempo, p, presión de un neumático, etc. Estas letras reemplazan a números. Así, cuando la velocidad es de 50 kilómetros por hora, podemos decir que $v = 50$; que $v = 25$, cuando es de 25 kilómetros, y etc.

En el caso del tiempo, ordinariamente lo medimos desde algún instante dado, que a menudo se denomina *tiempo de origen* o *tiempo cero*. Por ejemplo, el instante en que dejamos caer una pelota desde lo alto de un edificio es el tiempo cero, o, $t = 0$. Después que han transcurrido 3 segundos, decimos que $t = 3$, después de 5 segundos, $t = 5$, etc.

Muy a menudo, en la práctica, encontramos que una variable depende de algún modo de otra variable. Por ejemplo, la distancia que un cohete, o una pelota, recorre depende del *tiempo* del trayecto. Llamamos a la distancia *variable dependiente*. El tiempo se denomina *variable independiente*.

En general, el tiempo se considera como una variable independiente, mientras que cualquier variable que dependa de él es una variable dependiente. Así, el costo de la vida, que depende del tiempo, es una variable dependiente. La temperatura externa también lo es.

Pueden encontrarse otras variables independientes además del tiempo. Por ejemplo, el área de un círculo depende del radio. Podemos llamar al área A variable dependiente, y al radio r, variable independiente.

¿QUÉ ES UNA FUNCIÓN?

Si una variable depende de otra, los matemáticos dicen que la primera variable es una

función de la segunda. La distancia que un cohete o una pelota recorren es una función del tiempo de recorrido. El área de un círculo es una función del radio del mismo.

Supongamos que designamos con la letra x cualquier variable independiente tal como el tiempo, el radio de un círculo, etc. Supongamos, además, que designamos con la letra y cualquier variable dependiente de x, tal como la distancia recorrida por un cohete, o el área de un círculo. En tal caso los matemáticos abrevian a menudo el enunciado de que *y depende de x* o que *y es una función de x*, escribiendo $y = f(x)$. Se lee *y es igual a f de x*. La letra f se emplea como una abreviatura de *función*.

Según esta notación funcional, podríamos abreviar la proposición de que el costo de la vida C es una función del tiempo t, escribiendo $C = f(t)$. Del mismo modo podríamos expresar la propiedad de que el área A de un círculo es una función del radio r, escribiendo $A = f(r)$.

Una función indica que hay una relación entre las variables dependiente e independiente. Así $C = f(t)$ indica que hay una relación entre el costo de vida C y el tiempo t. Del mismo modo, $A = f(r)$ indica que hay una relación entre el área de un círculo A y su radio r. Uno de los problemas importantes en matemáticas y sus aplicaciones a otros campos, tales como ciencia, ingeniería y economía, es determinar la naturaleza de las relaciones entre variables.

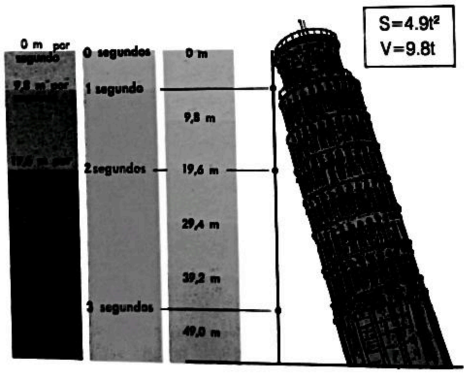
En algunos casos, la relación entre variables es sencilla. Por ejemplo, sabemos por la geometría que el área A de un círculo se expresa según su radio por medio de la relación $A = \pi r^2$, donde π es una constante cuyo valor es 3,14159 . . . , o aproximadamente 22/7.

En otros casos, la relación entre variables es muy difícil de obtener. Por ejemplo, la relación entre el costo de la vida y el tiempo no se conoce, pese a que podemos tener alguna idea acerca de tal relación basada en experiencias anteriores. De todos modos, aun cuando no podamos encontrar tal relación, no significa esto que no exista ninguna.

PROPORCIONES DE VARIACIÓN

Supongamos que a las 10:00 hs. un conductor de automóvil está a 30 kilómetros de determinada ciudad, y 2 horas después a 160 kilómetros de distancia. La variación de la distancia en 2 horas es entonces 160 — 30, o sea 130 kilómetros. Dividiendo esta variación de distancia por la variación de tiempo –o sea 130 kilómetros dividido por 2 horas– obtenemos 65 kilómetros por hora. Ésta es la *velocidad media* del conductor. Debería destacarse que, durante las 2 horas, el conductor pudo haber viajado a 80 kilómetros por hora parte del tiempo, a 50 kilómetros en otros momentos, o pudo haberse detenido algún rato. En *promedio*, de

VELOCIDAD' TIEMPO' DISTANCIA'



Supongamos que una pelota se arroja desde lo alto de la Torre Inclinada de Pisa. El diagrama a la izquierda muestra la velocidad de la pelota y la distancia que ha caído en tiempos determinados. Usando el análisis matemático se puede averiguar la velocidad y la distancia recorrida en cualquier instante dado.

todos modos, viajó a 65 kilómetros por hora. Vemos que la *velocidad media* es la razón entre el tiempo y la distancia recorrida. La velocidad real del conductor en un instante particular se denomina *velocidad instantánea*. Podemos tener una buena idea de la velocidad instantánea de un conductor hallando su velocidad promedio en un intervalo de tiempo muy breve. Supongamos, por ejemplo, que sabemos que a las 10:00 hs. el conductor está a 50 kilómetros de distancia de determinada ciudad y que 5 minutos después está a 55 kilómetros de distancia. Por tanto, ha recorrido 5 kilómetros en 5 minutos, o aproximadamente 1 kilómetro en 1 minuto, o sea 60 kilómetros por hora. Ésta es su *velocidad media*, pero es también una muy buena aproximación a su *velocidad instantánea*, pues si se detuvo durante ese tiempo o estuvo viajando mucho más lenta o rápidamente que a 60 kilómetros por hora, seguramente no pudo haberlo hecho durante un largo lapso.

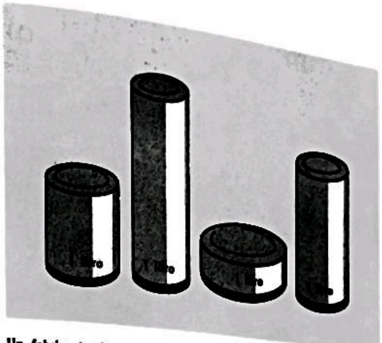
La velocidad media es, como hemos visto, la razón entre el tiempo y la distancia recorrida. La velocidad instantánea es la razón *instantánea* de cambio de distancia respecto del tiempo. El problema de hallar la razón instantánea de variación de distancia con respecto al tiempo, o de otras cantidades, es una de las partes más importantes del análisis matemático y tiene muchas aplicaciones. Trataremos de ver cómo pueden hallarse estas razones instantáneas de variación.

PROCEDIMIENTO DE DIFERENCIACIÓN

Consideremos una vez más una pelota que se deja caer desde un edificio alto. Si llamamos s a la distancia (en metros) que la pelota recorre en un tiempo t (en segundos), entonces s dependerá de t; o sea, s será una función de t, o sea $s = f(t)$. Una fórmula que expresa esa relación entre s y t en el caso de que la resistencia del aire sea despreciable es ésta:

$$s = 4{,}9\,t^2 \qquad (1)$$

A través de esta fórmula vemos que en $t = 0$ tenemos $s = 0$, lo cual no es sorprendente, pues la pelota cae una distancia cero en un tiempo cero. Después de 1 segundo –o sea, $t = 1$– vemos por la ecuación (1) que $s = 4{,}9$, lo que significa que en 1 segundo la pelota ha caído 4,9 metros. Del mismo modo, después



Un fabricante desea hacer recipientes de un litro. Usando el análisis matemático puede determinar el área superficial mínima necesaria para contener el volumen requerido.

de 2 segundos –o sea, $t = 2$– vemos que $s = 19{,}6$; por tanto, en 2 segundos la pelota ha caído 19,6 metros.

Supongamos ahora que queremos hallar la velocidad de la pelota en el tiempo t, o sea, la velocidad instantánea en el tiempo t. Incrementamos el tiempo t en una pequeña cantidad, que simbolizamos por dt. Podemos pensar a dt como una "pequeña parte de t".

Tratamos ahora de hallar la distancia que la pelota recorrerá en el tiempo $t + dt$, o sea, el tiempo t más una pequeña fracción extra de tiempo dt. Esta distancia será la distancia original s más una pequeña parte extra de distancia, que denominamos ds. Puesto que la distancia recorrida en un tiempo $t + dt$ es $s + ds$, deducimos de la fórmula (1) que

$$s + ds = 4{,}9(t + dt)^2 \qquad (2)$$

que puede escribirse

$$s + ds = 4{,}9\,t^2 + 9{,}8\,t(dt) + 4{,}9(dt)^2 \qquad (3)$$

La pequeña parte extra de distancia ds que la pelota recorre se obtiene restando s del lado izquierdo de (3) y su valor equivalente $4{,}9\,t^2$ del lado derecho de (3). De este modo hallamos

$$ds = 9{,}8\,t(dt) + 4{,}9(dt)^2 \qquad (4)$$

Ahora bien, como dt representa un número pequeño, $(dt)^2$ es un número muy pequeño; tan pequeño, en realidad, que el último término de la derecha de (4) puede eliminarse en la práctica. Así tenemos

$$ds = 9{,}8\,t(dt) \qquad (5)$$

Dividiendo ambos miembros por dt, obtenemos

$$\frac{ds}{dt} = 9{,}8\,t \qquad (6)$$

La cantidad a la izquierda de (6), que es la pequeña parte de distancia ds dividida por la pequeña parte de tiempo dt, es la velocidad instantánea de la pelota en el tiempo t. Si designamos con v la velocidad instantánea, tenemos

$$v = 9{,}8\,t \qquad (7)$$

Si reemplazamos $t = 3$ en (7) hallamos que $v = 29{,}4$. Esto significa que, después de 3 segundos, la pelota está viajando a 29,4 metros por segundo, que es su velocidad instantánea.

El procedimiento que hemos utilizado para obtener el resultado (6) se conoce en análisis matemático como *diferenciación*. La parte del análisis que trata de dichos procedimientos se denomina cálculo diferencial. La cantidad ds/dt se denomina *derivada de s respecto de t*, o simplemente *derivada de s*.

Si queremos, podemos utilizar la (7) para hallar la derivada de v respecto de t. Para hacerlo, usamos exactamente el mismo procedimiento dado antes. Incrementamos el tiempo t en dt de modo que el tiempo total es $t + dt$. Por tanto, la velocidad v aumenta en una pequeña parte, que denominamos dv; entonces, la nueva velocidad es $v + dv$. De (7) deducimos que

$$v + dv = 9{,}8(t + dt) \qquad (8)$$

o

$$v + dv = 9{,}8\,t + 9{,}8\,dt \qquad (9)$$

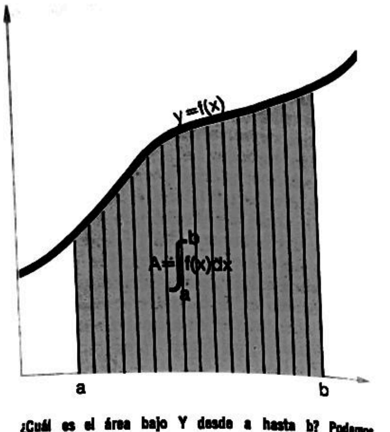
Restando v del lado izquierdo de (9) y su valor equivalente 9,8 t del lado derecho, tenemos

$$dv = 9{,}8\,dt \qquad (10)$$

o sea

$$\frac{dv}{dt} = 9{,}8 \qquad (11)$$

El lado izquierdo de (11) es la razón instantánea de cambio de velocidad respecto del tiempo, que se denomina *aceleración instantánea*. De (11) vemos que la aceleración instantánea es una constante. El significado es que cuando la pelota cae hacia la Tierra, aumenta su velocidad en 9,8 metros por segundo en cada segundo.



¿Cuál es el área bajo Y desde a hasta b? Podemos dividir el área en regiones casi rectangulares cuyas áreas pueden hallarse aproximadamente. La suma de esas áreas está muy próxima a la que queremos. La integración da la "mayor aproximación posible", que es el área real bajo la curva.

Dado que v es la derivada de s, deducimos que dv/dt es la derivada de la derivada de s, que se denomina la *derivada segunda* de s. Se escribe como:

$$\frac{d^2s}{dt^2} = 9{,}8 \qquad (12)$$

APLICACIONES

Como ya hemos mencionado, el procedimiento de diferenciación –o sea, encontrar derivadas– se estudia en la parte del análisis matemático conocida como *cálculo diferencial*. Hay muchas aplicaciones del cálculo diferencial además de las que implican velocidad y aceleración.

Una aplicación muy importante del cálculo diferencial es la de los *máximos* y los *mínimos*.

Para ver qué implican tales aplicaciones, consideremos un problema particular. Supongamos que fabricamos envases de metal para una compañía de sopas. La compañía nos pide que fabriquemos latas de forma cilíndrica, con la exigencia de que tengan una capacidad de un litro.

Podríamos hacer las latas altas y delgadas, o cortas y gruesas. ¿Cómo decidiremos qué hacer? Sabemos que el metal con que se hacen los envases costará dinero. Entonces es natural que nos preguntemos si podemos hacer el envasé requerido usando la menor cantidad –o sea, la cantidad mínima– de metal para la superficie total de la lata. Por el método del cálculo diferencial podemos determinar las medidas exactas, el diámetro y la altura, que debe tener el envase para contener un volumen dado y al mismo tiempo tener la menor área superficial posible, y por lo tanto el menor costo. Como estamos interesados en la menor, o mínima, área superficial, éste se llama un problema de hallar *mínimos*.

Como otro ejemplo, supongamos que un hombre tiene una plancha de cartón rectangular que mide 3 por 5 metros. Desea armar una caja abierta con ella cortando cuadrados iguales de las esquinas y luego doblando los bordes. La pregunta es: ¿de qué tamaño deberían cortarse los cuadrados para que la caja contenga el mayor volumen? Este problema, que consiste en hallar *máximos*, puede también resolverse por el cálculo diferencial.

EL PROCEDIMIENTO DE INTEGRACIÓN

Ya hemos visto cómo, si tenemos $s = 4{,}9\,t^2$, podemos hallar $ds = 9{,}8\,t(dt)$. El procedimiento consiste en comenzar con la distancia total s y hallar la pequeña parte de distancia ds recorrida en un tiempo dt. Como se recordará, dimos a este procedimiento el nombre de diferenciación.

Preguntamos ahora: si se nos da la ecuación $ds = 9{,}8\,t(dt)$, ¿podemos volver a $s = 4{,}9\,t^2$? Esto es lo opuesto, o inverso de la diferenciación. Como ds representa una pequeña parte de distancia, es natural que para hallar la distancia total s debamos sumar todas las pequeñas partes de distancia ds. El proceso de sumar, o adicionar, todas las pequeñas partes de distancia ds para obtener s se representa matemáticamente escribiendo

$$\int ds = s \qquad (13)$$

El símbolo $\int$ se denomina signo *integral*, y (13) se lee *la integral de ds es s*. El procedimiento de hallar integrales se denomina *integración* y se estudia en una parte del análisis matemático que se conoce como *cálculo integral*. Por métodos de cálculo integral, podemos hallar, por ejemplo, que

$$s = \int ds = \int 9{,}8\,t(dt) = 4{,}9\,t^2 \qquad (14)$$

por tanto, hemos recuperado la fórmula $s = 4{,}9\,t^2$. Se desprende que la integración es lo opuesto, o inverso, de la diferenciación.

APLICACIONES

Así como hay muchas aplicaciones del cálculo diferencial hay también muchas aplicaciones del cálculo integral.

Una aplicación importante es la de hallar áreas limitadas por curvas cerradas complicadas o la de hallar volúmenes limitados por superficies cerradas también complicadas. Otra aplicación es la de hallar la longitud total de una curva complicada o el área total de una superficie complicada. La idea comprendida en tales casos es la de suma, o adición, de pequeñas partes de área o de pequeñas partes de volumen, y entonces implica integración.

Hemos visto cómo el proceso de integración nos permite ir de la ecuación (6) nuevamente a la ecuación (1). La ecuación (6) es una ecuación que incluye una razón instantánea de cambio. Es una derivada. Las ecuaciones que incluyen derivadas de cantidades que queremos determinar se denominan *ecuaciones diferenciales*. Dichas ecuaciones a menudo aparecen en ciencia e ingeniería porque en muchos casos es más fácil llegar a una relación entre las derivadas de las cantidades que entre las cantidades mismas. El procedimiento de resolver ecuaciones diferenciales se conoce también a menudo como *integración* de ecuaciones diferenciales.

En los párrafos anteriores sólo hemos podido echar una ojeada a algunos de los muchos conceptos importantes del análisis matemático. Con el fin de apreciar el poder del análisis para resolver problemas importantes de las matemáticas, la ciencia y la ingeniería, el estudiante podrá consultar algunos de los muchos libros de análisis matemático disponibles.

TEORÍA DE LOS CONJUNTOS

Una manada de vacas, una bandada de pájaros, un cardumen de peces... Cada una de las palabras "manada", "bandada" y "cardumen" podría reemplazarse por la palabra "conjunto". Un *conjunto* es simplemente una colección de objetos o ideas.

El concepto de conjuntos fue desarrollado como un nuevo ramo de las matemáticas, a fines del siglo XIX, por el matemático alemán Georg Cantor. Desde comienzos del siglo XX, la teoría de los conjuntos se ha desarrollado rápidamente, y hoy tiene importantes aplicaciones en prácticamente todos los ramos de las matemáticas. En verdad, la mayor parte de nuestras matemáticas puede derivarse de la teoría de los conjuntos.

Dos clases de notaciones para conjuntos son de uso común: una es la notación por enumeración, en la cual escribimos $\{1, 3\}$ para el conjunto consistente en los números 1 y 3. {Roberts, Roye} indica el conjunto de dos primeros presidentes de Liberia.

La segunda es la notación por comprensión, en la que escribimos $\{x|x$ es un número entero$\}$ para el conjunto de los enteros, y $\{y|y$ fue uno de los dos primeros presidentes de Liberia$\}$ para el conjunto {Roberts, Roye}. Leemos $\{x|x\ ...\}$ como "el conjunto de todas las x tales que x. ...". Así, $\{x|x$ es entero$\}$ se lee "el conjunto de todas las x tales que x es un entero".

De las dos notaciones, aquélla por comprensión es la más frecuentemente empleada en matemáticas. Obsérvese que la letra que se usa en ella no tiene significado especial; así, por ejemplo, $\{x|x$ es un entero$\} = \{y|y$ es un entero$\}$.

Usualmente se emplean mayúsculas para simbolizar los conjuntos. Por ejemplo, escribimos $A = \{x|x$ es un entero$\}$.

Los *miembros*, o *elementos*, de un conjunto son simplemente aquellas cosas que constituyen el conjunto. Así, los elementos del conjunto $\{2, 3\}$ son los números 2 y 3. Los elementos del conjunto $\{x|x$ es ciudadano de los Estados Unidos$\}$ son los ciudadanos de los Estados Unidos. Cuando decimos que dos conjuntos A y B son *iguales* (que se escribe $A = B$), queremos decir que cada elemento de A lo es de B, y viceversa, que todo elemento de B lo es de A. Por ejemplo, $\{5, 7\} = \{7, 5\}$. Obsérvese que no reviste significación el orden de enumeración de los elementos. Si x es elemento del conjunto A, escribimos $x \in A$. Si x no es elemento de A, escribimos $x \notin A$: así $2 \in \{2, 3\}$, pero $4 \notin \{2, 3\}$.

SUBCONJUNTOS DE CONJUNTOS

Cada elemento del conjunto $\{a, b\}$ es elemento del conjunto $\{a, b, c\}$. Decimos que $\{a, b\}$ es un *subconjunto* del $\{a, b, c\}$ y escribimos $\{a, b\} \subseteq \{a, b, c\}$. Cada elemento del conjunto $A = \{x|x$ es par$\}$, lo es del $B = \{x|x$ es un entero$\}$. Decimos que A es un subconjunto de B, y escribimos $A \subseteq B$.

Nuestros ejemplos ilustran el uso "natural" de la palabra "subconjunto" para indicar una parte de un todo (como, por ejemplo, *subtotal* es una parte de un total). Empero, resulta útil considerar que $\{a, b\} \subseteq \{a, b\}$. En general, cada conjunto es un subconjunto de sí mismo: $A \subseteq A$ para todos los conjuntos A.

Si A y B son conjunto, $A \subseteq B$ significa que, siempre que $x \in A$, entonces $x \in B$. Si A es un subconjunto de B, pero $A \neq B$, escribimos $A \subset B$, y decimos que A es subconjunto *propio* de B. Así, por ejemplo, $\{a, b\}$ es un subconjunto de $\{a, b\}$, pero no un subconjunto propio de $\{a, b\}$, mientras que $\{a, b\}$ es al mismo tiempo un subconjunto y un subconjunto propio de $\{a, b, c\}$.

UNIÓN DE CONJUNTOS

Así como dos números pueden combinarse por suma o productos para dar un tercero, del mismo modo dos conjuntos pueden combinarse de varias maneras para dar un tercero. En particular: dados dos conjuntos A y B, podemos formar su *unión* $A \cup B$. Éste es el conjunto cuyos elementos lo son de A o B (o de ambos). Así, por ejemplo:

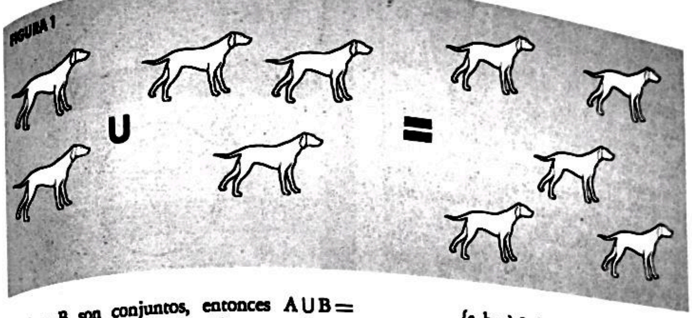
$$\{a, b\} \cup \{c, d\} = \{a, b, c, d\}$$

y

$$\{x|x \text{ es impar}\} \cup \{x|x \text{ es par}\} = \{x|x \text{ es un entero}\}.$$

En el "lenguaje de la teoría de los conjuntos", si A y B son conjuntos, entonces $A \cup B = \{x|x \in A, \text{ o } x \in B \text{ (o ambos)}\}$.

FIGURA 1



Este concepto de unión se utiliza en muchas partes de las matemáticas. En la escuela primaria, por ejemplo, se da a los niños la idea de la suma de números enteros por medio de la unión de conjuntos. Así, la unión de los dos conjuntos que se muestran en la Figura 1 corresponde al hecho aditivo $2 + 3 = 5$. En geometría, el concepto de unión se emplea para definir un triángulo como la unión de tres segmentos, cada uno de los cuales tiene un punto extremo común con los otros dos, tal como se puede ver en la Figura 2.

Supongamos que se considera:

$$\{a, b\} \cup \{a, c\} = \{a, b, c\}.$$

¿Podemos escribir también

$$\{a, b\} \cup \{a, c\} = \{a, a, b, c\}?$$

Sí, podemos; es cierto que $\{a, b, c\} = \{a, a, b, c\}$, pero no "sumamos" al conjunto $\{a, b\}$ al repetir el símbolo "a". Por ejemplo, considérese el conjunto {usted, lector de este artículo} = {usted} = {lector de este artículo}. Usted no puede volverse dos personas por nombrarse a sí mismo dos veces. Luego, al enumerar los elementos de un conjunto, coincidimos en no incluir un elemento más de una vez.

INTERSECCIÓN DE CONJUNTOS

Similar al concepto de unión de dos conjuntos es el de *intersección* de dos conjuntos. Si A y B son conjuntos, la intersección de A y B, $A \cap B$, es

$$A \cap B = \{x|x \in A \text{ y } x \in B\}.$$

Así, por ejemplo

$$\{a, b, c\} \cap \{a, e, f\} = \{a\}$$

y

$$\{x|x \text{ es un ser humano de ojos castaños}\} \cap \{x|x \text{ es una mujer}\} = \{x|x \text{ es una mujer de ojos castaños}\}.$$

Al igual que el de unión de conjuntos, el concepto de su intersección encuentra amplia aplicación en matemáticas. Por ejemplo: el conjunto de divisores comunes de 12 y 18,

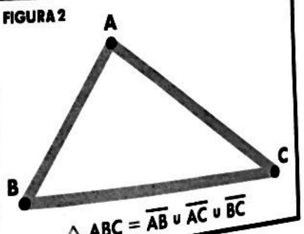
$$\{x|x \text{ divide a } 12\} \cap \{x \text{ divide a } 18\} = \{1, 2, 3, 4, 6, 12\} \cap \{1, 2, 3, 6, 9, 18\} = \{1, 2, 3, 6\}.$$

En geometría, podemos simbolizar el hecho de que dos rectas L_1 y L_2 se intersecan en un punto p, por

$$L_1 \cap L_2 = p.$$

¿Qué pasa con $\{x|x$ es par$\} \cap \{x|x$ es impar$\}$? Puesto que ningún número es al mismo tiempo par e impar, puede parecer que no hay respuesta posible. Sin embargo, para tratar ésta y muchas otras situaciones similares, introducimos el concepto de conjunto *vacío*. También llamado conjunto *nulo*, el vacío se simboliza por { }, o por $\varnothing$. Así:

FIGURA 2



$\triangle ABC = \overline{AB} \cup \overline{AC} \cup \overline{BC}$

$$\{x|x \text{ es par}\} \cap \{x|x \text{ es impar}\} = \emptyset.$$

Análogamente, si las rectas L_1 y L_2 son paralelas, $L_1 \cap L_2 = \emptyset$.

Con frecuencia es útil considerar un conjunto fijo, llamado conjunto *universal*, tal que todos los considerados en un momento dado son subconjuntos del universal. Por ejemplo, si estamos discutiendo subconjuntos de los enteros, tales como el conjunto de los naturales {1, 2, 3, 4, ...} y el de los pares {0, 2, 4, ...}, podríamos tomar nuestro conjunto universal como el conjunto de los números enteros. De igual modo tomaríamos nuestro conjunto universal como el de todos los puntos del plano si estuviésemos discutiendo subconjuntos del conjunto de puntos de un plano, tales como los conjuntos de puntos que forman círculos y los de los que constituyen triángulos.

COMPLEMENTOS DE CONJUNTOS

Supongamos que nuestro conjunto universal es el de todos los alumnos del Colegio Secundario de Podunk, y que A es el conjunto de todas las alumnas de dicho Colegio; luego, al conjunto de todos los alumnos varones de la Secundaria de Podunk se lo llama *complemento* del conjunto A, y se lo indica por A′. El complemento A′ de un conjunto A consiste en todos los elementos del universal que no sean elementos de A. En símbolos, si $A \subseteq U$, entonces

$$A' = \{x|x \in U \text{ y } x \notin A\}.$$

De este modo, si nuestro conjunto universal es el de los números reales, y A es el de los racionales, entonces A′ es el conjunto de los irracionales.

DIAGRAMAS DE VENN

A menudo, es provechoso representar conjuntos, y sus relaciones, mediante lo que se denomina comúnmente *diagramas de Venn.* En la Figura 3, la región encerrada por un rectángulo representa el conjunto universal, y las que están dentro del rectángulo, rodeadas por círculos (u otras curvas), representan subconjuntos del universal. La Figura 4 muestra diagramas de Venn que ilustran algunos conceptos de la teoría de los conjuntos.

Los diagramas de Venn pueden utilizarse para volver plausibles varios enunciados (si bien no para probarlos en realidad) de igualdad

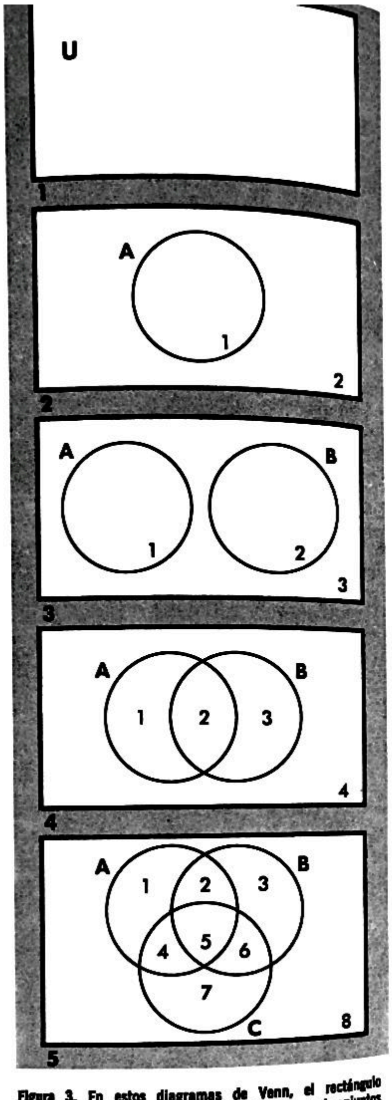


Figura 3. En estos diagramas de Venn, el rectángulo representa el conjunto universal, U. Los subconjuntos se representan por círculos y están marcados A, B y C; estos subconjuntos dividen al U en regiones señaladas con números. En el diagrama 3, los conjuntos A y B son "disjuntos". En el diagrama 4, el conjunto A consiste en las regiones 1 y 2; el B, en las 2 y 3. En el diagrama 5, A consiste en las regiones 1, 2, 4 y 5.

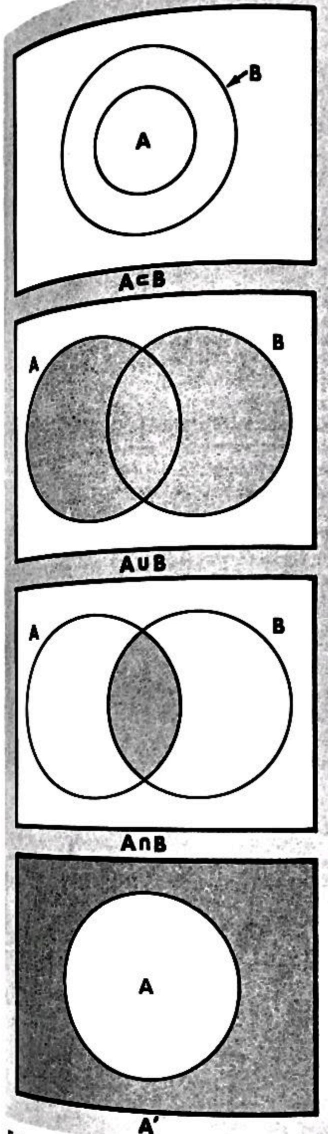
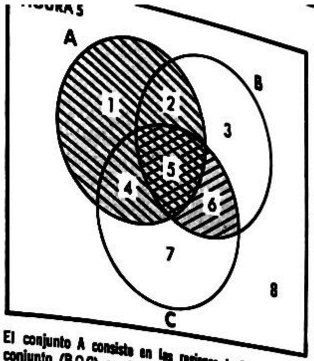


Figura 4. En el diagrama superior, A es un subconjunto propio de B. En los tres de abajo, las zonas sombreadas representan los conjuntos indicados en cada diagrama.



El conjunto A consiste en las regiones 1, 2, 4 y 5. El conjunto $(B \cap C)$ consiste en las regiones 5 y 6. ¿Cuáles regiones integran el conjunto $A \cup (B \cap C)$?

entre conjuntos, tales como las dos igualdades que en álgebra se denominan propiedades distributivas:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

y

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

(Compárese $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ del álgebra ordinaria.) Así, la región sombreada de la Figura 5 ilustra $A \cup (B \cap C)$, mientras que la doblemente sombreada de la Figura 6 lo hace con $(A \cup B) \cap (A \cup C)$; en ambos casos obtenemos la misma región.

En la Figura 7 se indican las situaciones correspondientes a $A \cap (B \cup C)$ y $(A \cap B) \cup (A \cap C)$.

ÁLGEBRA DE CONJUNTOS

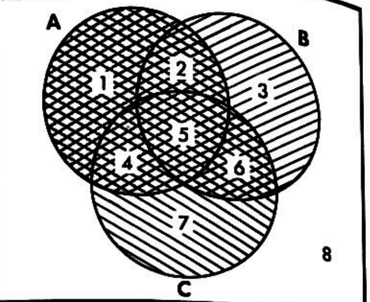
Existen muchas analogías entre las operaciones de unión e intersección de conjuntos y las operaciones de suma y multiplicación de números. Es fácil ver que, por ejemplo, así como es $a + b = b + a$ y $a \cdot b = b \cdot a$ para todos los números a y b, es también cierto que, para todos los conjuntos A y B, $A \cup B = B \cup A$, y $A \cap B = B \cap A$. Análogamente, así como 0 tiene la propiedad de que $a + 0 = a$, y $a \cdot 0 = 0$ para todo número a, asimismo $A \cup \emptyset = A$ y $A \cap \emptyset = \emptyset$ para todo conjunto A. Además, del mismo modo que $a \cdot 1 = a$ para todo número a, es $A \cap U = A$ para todo conjunto $A \subseteq U$ (conjunto universal).

Empero, las analogías no son completas; existen propiedades de nuestro sistema numérico para las cuales no hay propiedades análogas en el álgebra de conjuntos, y hay propiedades de ésta que no tiene otras análogas en nuestro sistema de números; por ejemplo: dado un nú-

mero a distinto de cero, existe un número b (el inverso de a), tal que $a \cdot b = 1$. Ahora bien, U es análogo a 1 en el sentido de que $A \cap U = A$ para todo conjunto A y, aun así, a menos que $A = U$, no existe conjunto B tal como $A \cap B = U$.

Por otro lado, el hecho de que $A \cup A = A \cap A = A$ para todos los conjuntos A, carece de su análogo en aritmética. Hemos establecido que $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ para todos los conjuntos A, B y C, y esto sí tiene analogía en $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. No obstante, también hemos dicho que $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ para todos los conjuntos A, B y C, y esto corresponde a $a + (b \cdot c) = (a + b) \cdot (a + c)$ para todos los números a, b y c: sin embargo, el último enunciado no es cierto. (Inténtese con $a = 1$, $b = 2$, y $c = 3$, por ejemplo.)

FIGURA 6



Arriba se indican los conjuntos (A∪B) y (A∪C) por líneas oblicuas; su intersección es la región en la que las rayas se entrecruzan: es el mismo conjunto que el de la Figura 5.

En la tabla siguiente, damos una lista de varias identidades que son ciertas para todos los conjuntos A, B y C que sean subconjuntos de alguno universal U: obsérvese el paralelismo entre las identidades de las columnas de derecha e izquierda.

(1) $A \cup A = A$	(1′) $A \cap A = A$
(2) $A \cup B = B \cup A$	(2′) $A \cap B = B \cap A$
(3) $A \cup (B \cup C) = (A \cup B) \cup C$	(3′) $A \cap (B \cap C) = (A \cap B) \cap C$
(4) $A \cup (B \cup C) = (A \cup B) \cap (A \cup C)$	(4′) $A \cap (B \cap C) = (A \cap B) \cup (A \cap C)$
(5) $A \cup \varnothing = A$	(5′) $A \cap U = A$
(6) $A \cup A' = U$	(6′) $A \cap A' = \varnothing$
(7) $U' = \varnothing$	(7′) $\varnothing' = U$
(8) $A \cup U = U$	(8′) $A \cap \varnothing = \varnothing$
	(9) $(A')' = A$

El área sombreada de la Figura 7A representa el conjunto A∩(B∪C), mientras que la superficie sombreada de la Figura 7B representa al conjunto (A ∩ B) ∪ (A ∩ C). Ambos conjuntos son el mismo, lo que confirma la identidad.

FIGURA 7A

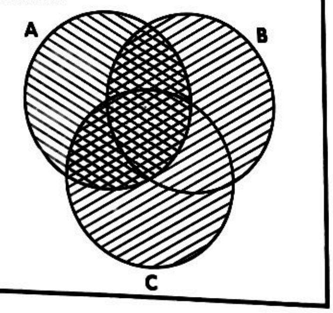
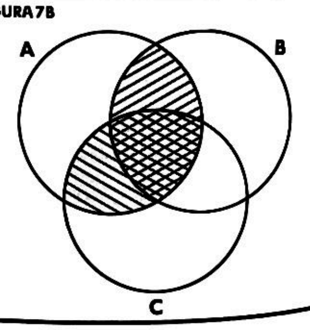


FIGURA 7B



Se ve con bastante facilidad que todas estas identidades son ciertas, como consecuencia de las definiciones de unión, intersección, conjunto vacío, conjunto universal y complemento, y, tal como hemos visto, pueden hacerse plausibles mediante el empleo de diagramas de Venn.

INFINITO Y FINITO

¿Qué queremos decir con "infinito"? Mucha gente piensa en "infinito" simplemente como en un número muy grande. Sin embargo, cuando un matemático habla de infinito, no piensa en un conjunto con un gran número de elementos –tal como el conjunto de granos de arena de la playa–; en verdad, encuentra que vale la pena evitar por completo cualquier referencia directa al recuento, al describir la diferencia entre conjuntos finitos e infinitos.

Para hacer esto, considera primero el concepto de conjuntos *correspondientes* o, en lenguaje más técnico los que se hallan en *correspondencia uno a uno* (1 — 1): concepto éste muy natural aún en etapas tempranas de instrucción matemática, en que el niño relaciona cada cubo sobre una mesa con un dedo de la mano.

La que sigue es una definición formal:

Una correspondencia 1 — 1 entre dos conjuntos A y B es el apareamiento de los elementos de A con los de B, de modo y forma tales que cada elemento de A tenga por par precisamente un elemento de B, y que cada elemento de B quede apareado con un determinado elemento de A.

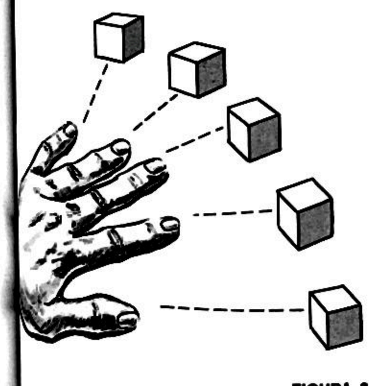


FIGURA 8

Si existe tal correspondencia de 1 — 1 entre los dos conjuntos no vacíos A y B, decimos que los dos conjuntos son *equivalentes*, y escribimos $A \sim B$. (También decimos que $\varnothing \sim \varnothing$, aunque no usaremos aquí este hecho.)

Ahora bien, para los conjuntos finitos, normalmente no nos molestamos en intentar la formación de pares para ver si ambos son equivalentes: simplemente, contamos los elementos de cada uno. Así, observamos que cada uno de los dos conjuntos que se muestran en la Figura 8 tiene cinco elementos, de modo que "sabemos" que son equivalentes. Sin embargo, en realidad el concepto de correspondencia 1 — 1 precede a la idea del recuento: por ejemplo, "cinco" es simplemente el nombre del número que le adjudicamos a los conjuntos que guardan correspondencia 1 — 1 con los dedos de una mano. Mucho antes de que supiese contar, el hombre primitivo podía vigilar el tamaño de su rebaño equiparando cada animal con un guijarro: si dejaba caer uno en un recipiente a medida que cada animal abandonaba el redil por la mañana, y luego lo quitaba a medida que cada animal retornaba al anochecer, sabía sin contar si volvía el mismo número de animales que había salido.

Ahora, si tenemos un conjunto finito A, sentimos intuitivamente que ningún subconjunto propio de A guarda con él una correspondencia 1 — 1. Inténtese, por ejemplo, equiparar el conjunto de dedos, menos el pulgar de una mano, con todo el conjunto de dedos de la otra mano. Pero, ¿qué pasa con los conjuntos infinitos? Considérese el conjunto $N = \{1, 2, 3, \ldots\}$ de los naturales, y el $E = \{2, 4, 6, \ldots\}$ de los naturales pares; claramente, E es subconjunto propio de N y, aun así, $E \sim N$, tal como se muestra por la correspondencia 1 — 1 de la Figura 9.

Generalizamos estos dos ejemplos para elaborar una definición formal:

Un conjunto *finito* es el que carece de subconjunto propio equivalente al mismo. Uno *infinito* es un conjunto que tiene por lo menos un subconjunto propio que equivale al mismo.

NÚMERO CARDINAL DE UN CONJUNTO INFINITO

Si se tienen conjuntos finitos equivalentes, éstos tienen el mismo número de elementos; es decir, son *equinumerosos*. ¿Qué ocurre con los

infinitos? Por ejemplo, ¿es equinumeroso el conjunto E de los números pares con el N de los naturales? Desde un punto de vista, es natural, ciertamente, argüir que E y N no son equinumerosos, dado que quitamos algo del conjunto N = {1, 2, 3, ...}, para obtener el E = {2, 4, 6, ...}. Si quitamos algo, digamos 1, del conjunto {1, 2, 3}, el conjunto que obtenemos {2, 3} no es equinumeroso con el {1, 2, 3}.

De todos modos, resulta útil aceptar que E y N son equinumerosos y, en general, dar la siguiente definición formal:

Se dice que dos conjuntos A y B (finitos o infinitos) son ***equinumerosos***, o que tienen el mismo ***número cardinal***, sólo si $A \sim B$.

Así, llegamos a la conclusión –a partir de esta definición– de que no sólo tienen el mismo número cardinal los conjuntos finitos de la Figura 8, sino que también lo poseen los conjuntos E y N de la Figura 9.

¿Tienen el mismo número cardinal todos los conjuntos finitos? Considérese, por ejemplo, el conjunto Q de los racionales positivos: el conjunto de todos los naturales junto con todos los fraccionarios positivos, $\frac{1}{2}$, $\frac{17}{19}$, $1\frac{7}{8}$, $\frac{1}{252}$, 0,12, etc. ¿Son equinumerosos Q y N? De nuevo presentimos intuitivamente que la respuesta debiera ser "no", ya que Q contiene muchos (una cantidad infinita de) números que no están en el conjunto N; sin embargo, podemos demostrar que $Q \sim N$, de modo que, según nuestra definición, Q y N son equinumerosos.

Un argumento simple de que $Q \sim N$ se puede basar en el diagrama que aparece en la Figura 10. Muestra los números racionales positivos en un ordenamiento que se extiende indefinidamente hacia la derecha y hacia abajo. (Por el momento, dejemos de lado las líneas.)

Supongamos ahora que se nos da un número $n \in N$, y queremos establecer una regla para asociar con n un número racional definido, $r \in Q$. Simplemente, seguimos la línea del diagrama.

A medida que lo hacemos, contamos 1, 2, 3, ..., n al recorrer cada fracción, excepto que no contamos las repeticiones. Por ejemplo: supongamos que tenemos $n = 5$. Luego, a lo largo de la línea, hallamos $\frac{1}{1}$, $\frac{2}{1}$, $\frac{1}{2}$, $\frac{1}{3}$, y $\frac{3}{1}$ (omitiendo $\frac{2}{2} = \frac{1}{1}$) y llegamos a la conclusión de que, correspondiendo con $5 \in N$, tenemos $\frac{3}{1} \in Q$. En forma similar, si $n = 13$, obtenemos $\frac{1}{1}$, $\frac{2}{1}$, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{3}{1}$, $\frac{4}{1}$, $\frac{3}{2}$, $\frac{2}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, $\frac{5}{1}$, $\frac{6}{1}$, y $\frac{5}{2}$ (omitiendo $\frac{2}{2} = \frac{1}{1}$, $\frac{2}{4} = \frac{1}{2}$, $\frac{3}{3} = \frac{1}{1}$, y $\frac{4}{2} = \frac{2}{1}$). Así, para $13 \in N$, corresponde $\frac{5}{2} \in Q$.

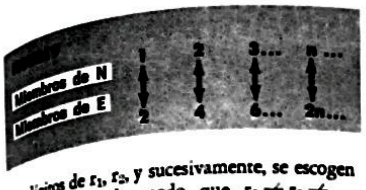
Recíprocamente: dado cualquier $r \in Q$, tomamos su representación como una fracción con los términos menores (tales como $\frac{1}{2}$ para $\frac{3}{6}$), y contamos "cuán lejos" está de $\frac{1}{1}$ a lo largo de la línea (omitiendo nuevamente las repeticiones). Así, si tenemos dados $\frac{4}{6} \in Q$, observamos que $\frac{4}{6} = \frac{2}{3}$, y contamos $\frac{1}{1}$, $\frac{2}{1}$, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{3}{1}$, $\frac{4}{1}$, $\frac{3}{2}$, $\frac{2}{3}$, de modo que, correspondiendo con $\frac{2}{3} \in Q$, tenemos $8 \in N$.

En este punto, podría suponerse que todos los conjuntos infinitos tienen el mismo número cardinal que N: ésta ciertamente es una conjetura plausible en esta etapa. Empero, resulta que, si sumamos al conjunto Q de los racionales positivos todos los irracionales positivos –tales como $\sqrt{2}$, $\sqrt[3]{2}$, π, o $\sqrt[3]{1 + \sqrt{2}}$, y así sucesivamente– obtenemos un conjunto con un cardinal diferente al de N. Inclusive, si consideramos solamente el conjunto R^* de los racionales e irracionales entre 0 y 1, obtenemos –tal como demostró Georg Cantor en 1874– un conjunto "mayor" que no puede coordinarse con N. La unión del conjunto de los racionales con el de los irracionales es el conjunto de los números reales.

La prueba de que R^* no es equinumeroso con N se basa en el hecho de que cada real comprendido entre 0 y 1 puede escribirse como decimal infinito. Por ejemplo: $\frac{1}{2} = ,5000 \ldots$, $\frac{1}{3} = ,333 \ldots$, $\frac{1}{7} = ,14285714857 \ldots$, $\sqrt{\frac{1}{2}} = ,7071 \ldots$, donde los racionales (tales como $\frac{1}{2}$, $\frac{1}{3}$, y $\frac{1}{7}$) se expresan como decimales periódicos, y los irracionales (tales como $\sqrt{\frac{1}{2}}$), como decimales no periódicos.

Supongamos ahora que sostenemos que tenemos correspondencia 1 — 1 entre los elementos de R^* y los de N, tal que:

$$
\begin{array}{cl}
N & \quad R^* \\
\downarrow & \\
1 \leftrightarrow & ,183478412001 \ldots = r_1 \\
2 \leftrightarrow & ,369715400000 \ldots = r_2 \\
3 \leftrightarrow & ,579321715432 \ldots = r_3 \\
4 \leftrightarrow & ,481762314000 \ldots = r_4 \\
5 \leftrightarrow & ,673216732167 \ldots = r_5 \\
6 \leftrightarrow & ,591416789143 \ldots = r_6 \\
7 \leftrightarrow & ,001326841841 \ldots = r_7 \\
. \leftrightarrow & , \ldots \ldots \quad \ldots = \ldots
\end{array}
$$



Los dígitos de r_1, r_2, y sucesivamente, se escogen arbitrariamente, de modo que $r_1 \neq r_2 \neq \ldots$

Ahora, formamos otros número en R^* del modo siguiente: como primer dígito a la derecha de la coma decimal se escoge cualquiera diferente del primero de r_1. Como segundo dígito se escoge cualquiera diferente del segundo de r_2, y siguiendo así. De esta manera, el número resultante podría ser ,2783254 ... ó ,3723689 ... Sin embargo, ninguno de estos números así seleccionados puede estar en nuestra lista, puesto que ,2783254... $\neq r_1$ porque difiere de éste en, por lo menos, el primer lugar decimal; ,2783254... $\neq r_2$ porque difiere de éste en el segundo lugar decimal, y así sucesivamente. En consecuencia, ,2783254... no estaba presente en la (supuestamente completa) lista y por eso nuestra correspondencia no es 1 — 1, como se suponía; no importa qué listado propongamos, el mismo proceso ha de demostrar que es incompleto, y arribamos a la conclusión de que no es posible ninguna correspondencia 1 — 1 entre R^* y N.

Los matemáticos indican el número cardinal de N por $\aleph_0$ (que se lee "alef nula": alef es la primera letra del alfabeto hebreo). Designan el número cardinal de los números reales con c. Una pregunta famosa en matemáticas es: ¿existe un conjunto infinito de números cuyo número cardinal esté entre $\aleph_0$ y c?; es decir, ¿existe un subconjunto S de los números reales R^* tal que el cardinal de S no sea ni $\aleph_0$ ni c? La ***hipótesis del continuo***, según Cantor, era la conjetura de que no existe tal conjunto S.

La cuestión de la verdad o falsedad de la hipótesis del continuo atrajo la atención de muchos matemáticos de primera línea, después de haber sido formulada por Cantor, pero no se la dirimió hasta 1963; empero, no se la resolvió con un simple "sí" o "no" (como esperaba la mayoría de los matemáticos); lo que finalmente se demostró en el trabajo de Kurt Gödel y Paul Cohen era que, dentro del marco de la teoría de los conjuntos admitida por la mayoría de los matemáticos actuales, la aceptación o el rechazo de la hipótesis del continuo lleva a sistemas de matemáticas igualmente válidos. Esto deja abierta todavía la cuestión de si existe, o no, otra formulación igualmente "útil" de la teoría de los conjuntos, en la que se pueda demostrar que la hipótesis del continuo es verdadera o falsa; en un sentido, pues, la hipótesis del continuo sólo se ha dirimido en forma relativa, y aún continúa la investigación de este problema y los que le están relacionados.

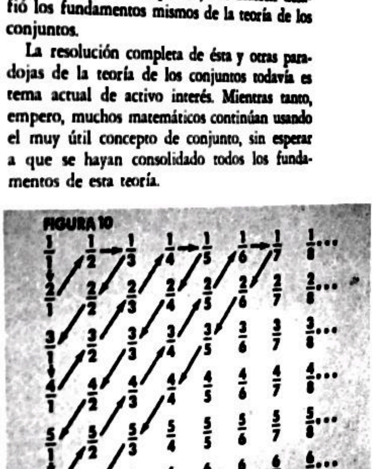
PARADOJAS DE LA TEORÍA DE LOS CONJUNTOS

Hay muchas paradojas asociadas a conjuntos que todavía interesan a los matemáticos de hoy en día. Una muy famosa fue la formulada por el renombrado matemático y filósofo Bertrand Russell (1872-1969), y se la conoce como la paradoja de Russell: imagínese un barbero en un pueblo; el barbero –dijo Russell– sólo afeita a los hombres que no lo hacen por sí mismos. La paradoja es: ¿quién afeita al barbero?

Si no lo hace por sí mismo, entonces se afeita a sí mismo, pero, si se afeita a sí mismo, no se puede afeitar por sí mismo. Cuando se lo expresó matemáticamente –en términos de clases de conjuntos– la paradoja de Russell desafió los fundamentos mismos de la teoría de los conjuntos.

La resolución completa de ésta y otras paradojas de la teoría de los conjuntos todavía es tema actual de activo interés. Mientras tanto, empero, muchos matemáticos continúan usando el muy útil concepto de conjunto, sin esperar a que se hayan consolidado todos los fundamentos de esta teoría.

FIGURA 10



NÚMEROS BINARIOS

El sistema binario numérico, o de numera-
ción, ofrece una forma interesante de abordar
el mundo de las matemáticas. Es el sistema más
simple en el que pueden realizarse la suma y
la multiplicación, porque hay sólo cuatro casos
de suma y cuatro de multiplicación que apren-
der. Por el contrario, en el sistema decimal
deben aprenderse 100 casos de suma y 100 de
multiplicación.

Sin embargo, la importancia del sistema binario en las matemáticas modernas reside en su aplicación práctica en la tecnología de la computación. Las computadoras son actualmente en su mayor parte binarias; o sea, usan números binarios para computar y procesar los datos. Con estos números, las computadoras modernas pueden efectuar hasta 1.000.000.000 de sumas por segundo.

Antes de discutir el sistema binario de numeración daremos una ojeada más detallada al sistema de numeración indoarábigo y al concepto de valor posicional en él. Comprender el valor decimal es fundamental para la comprensión del sistema numérico binario.

SISTEMA DE NUMERACIÓN INDOARÁBIGO

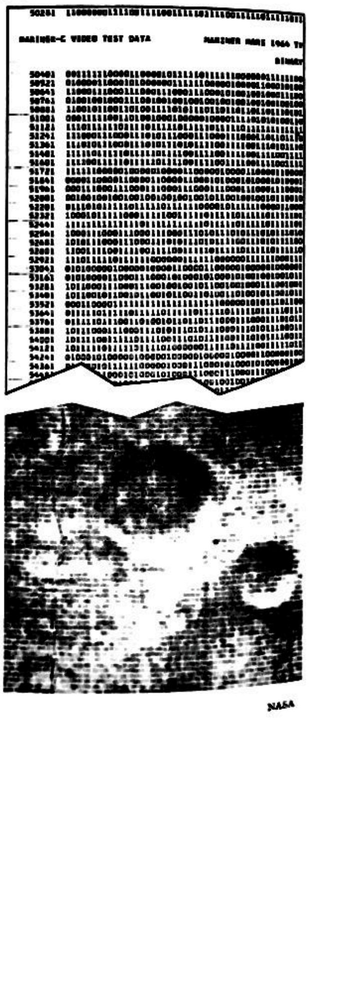
El sistema de numeración usado actualmente por el mundo civilizado se denomina sistema indoarábigo, probablemente porque lo concibieron los hindúes, y los árabes lo llevaron al mundo occidental. Es el único sistema de numeración que la mayoría de nosotros conoce, y se adapta muy bien a casi todas nuestras necesidades comerciales y técnicas.

El sistema indoarábigo se denomina a menudo *sistema decimal*, o *de base diez*, porque necesita sólo diez símbolos para representar cualquier número. Estos símbolos, que se denominan *dígitos*, son: 1, 2, 3, 4, 5, 6, 7, 8, 9 y 0.

El sistema de numeración binario se usa para trasmitir fotos del espacio a las estaciones receptoras en la Tierra. Las fotos se trasmiten como señales de radio, representando los ceros y unos del sistema binario (foto superior, derecha). La computadora convierte los dígitos en una serie de puntos. La foto de la superficie de Marte, a la derecha, se obtuvo en esta forma.

Estos diez símbolos representan los números uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve y cero, respectivamente.

El sistema indoarábigo se denomina también *sistema decimal posicional*, porque el número que representa cada dígito depende de su "posición" o "lugar" en el número. El último lugar de la derecha en cualquier número es la unidad, o posición uno. El valor del decimal de la siguiente posición a la izquierda es diez. En general, los lugares decimales de derecha, a izquierda, en cualquier número decimal, son: unidades, decenas, centenas, millares, etc. Por ejemplo, los valores decimales representados por 4.081 son:



NASA

$$4.081 = (4 \times 1.000) + (0 \times 100) + (8 \times 10) + (1 \times 1)$$

Otra forma de considerar el sistema decimal posicional es por medio de la idea de agrupamiento. Supongamos que tenemos un conjunto de 13 puntos marcados en una hoja de papel. Dibujamos un anillo alrededor de 10 de esos puntos. Quedarán 3 puntos. Tenemos un conjunto de 10 puntos más 3 puntos solos sobrantes. Este hecho puede expresarse como $13 = (1 \times 10) + (3 \times 1)$.

Supongamos ahora que tenemos 37 puntos y que los "cercamos" en grupos de 10. Tendremos 3 conjuntos de 10 puntos, más 7 puntos solos. Podríamos escribirlo como $37 = (3 \times 10) + (7 \times 1)$. Del mismo modo, si tuviéramos 128 puntos, podríamos tener 1 conjunto de cien ó 10 decenas, 2 conjuntos de 10, y 8 puntos restantes. Podríamos escribir $128 = (1 \times 100) + (2 \times 10) + (8 \times 1)$. Toda esta insistencia sobre las "decenas" nos parece bastante natural. Después de todo, tenemos diez dedos en las manos y diez en los pies, pero ¿cómo supone usted que estaría contando actualmente el hombre y qué tipo de sistema de numeración usaría, si no tuviera su anatomía estructurada como lo está?

SISTEMA DE NUMERACIÓN BINARIO

Imaginémonos a todos nosotros con un dedo en cada mano y un dedo en cada pie. Supongamos además que todos los números que usáramos pudieran expresarse con los "dígitos" 0 y 1. Como vamos a usar sólo dos símbolos para escribir cualquier número, podemos llamarlo *sistema binario*, o *de base dos*. En un sistema binario el valor de cualquier lugar en un número es dos veces mayor que el lugar a su derecha. Así, los valores posicionales –de derecha a izquierda– en un sistema binario son: unos, dos, cuatros, ochos, etc.

El número 1 escrito en el sistema binario se escribiría como 1_{dos}. En este caso, el número representado por 1_{diez} y por 1_{dos} es el mismo. El subíndice *dos* indica que estamos expresando números en el sistema binario; el subíndice diez indica el sistema decimal. ¿Cómo indicaríamos 2_{diez} como número binario? Tratemos de responder a esta pregunta.

Supongamos que tenemos un conjunto de 2 estrellas. Si dibujamos un anillo alrededor de

ellas, tenemos un conjunto de 2 estrellas, y
ninguna estrella sola sobrante. Escribiríamos
$2_{diez} = 10_{dos}$, que se lee como "1 dos más 0
unos". Si comenzamos con un conjunto de tres
estrellas y trazamos un anillo alrededor de dos
de ellas, tendríamos un conjunto de 2 estre-
llas más otro de 1 estrella. Esto se indicaría
como $3_{diez} = 11_{dos}$, leído como "1 dos más
1 uno".

¿Cómo se expresaría 4_{diez}? Observemos que
hemos estado haciendo pares de conjuntos equi-
valentes siempre que fuera posible. Así, si te-
nemos un conjunto de 4 estrellas, podríamos
primero formar 2 conjuntos de 2 estrellas cada
uno. Ahora tracemos un anillo alrededor de
estos conjuntos. Esta aproximación sugiere que
4_{diez} puede pensarse como 1 cuatro, ningún
dos y ningún uno, y escribirse como 100_{dos}.
De forma similar, $5_{diez} = 101_{dos}$, que es "1
cuatro, ningún dos, y 1 uno".

Veamos entonces cómo trataríamos al 6. Primero tendríamos 3 conjuntos de 2 estrellas cada uno. Luego aparearíamos 2 de esos conjuntos, finalizando con 1 conjunto de cuatro, 1 de dos y ningún uno. Escribimos 6_{diez} como 110_{dos}.

Usando el mismo desarrollo, se ve claramente que $7_{diez} = 111_{dos}$, interpretado como un conjunto de cuatro, uno de dos y 1 uno. Podemos

TABLA 1
Esquema del valor de posición

Binario				Decimal	
8	4	2	1	10	1
ochos	cuatros	doses	unos	decenas	unos
			0		0
			1		1
		1	0		2
		1	1		3
	1	0	0		4
	1	0	1		5
	1	1	0		6
	1	1	1		7
1	0	0	0		8
1	0	0	1		9
1	0	1	0	1	0
1	0	1	1	1	1
1	1	0	0	1	2
1	1	0	1	1	3
1	1	1	0	1	4
1	1	1	1	1	5

analizar 8_{diez} de la misma forma. Primero tenemos 4 conjuntos de 2 estrellas en cada uno. Luego tenemos 2 conjuntos con (2×2), o sea, 4 estrellas cada uno. Finalmente, tenemos un conjunto de $(2 \times 2 \times 2)$, o sea 8 estrellas en él. Esto se escribiría como 1000_{dos} y significaría 1 ocho, 0 cuatros, 0 dos, 0 unos.

Resumamos ahora lo que hemos aprendido acerca de la numeración binaria en la Tabla 1 y veamos si podemos expresar nuestras ideas.

¿Cómo expresaríamos 106_{diez} en base dos? Recordemos: los lugares decimales en la numeración binaria son unos, dos, cuatros, ochos, dieciséis, treinta y dos, sesenta y cuatros, ciento veintiochos, etc. El mayor valor decimal contenido en 106 es sesenta y cuatro. 106 contiene 1 sesenta y cuatro, y $106 - 64 = 42$. Hay 1 treinta y dos en 42; $42 - 32 = 10$. Hay 0 dieciséis en 10, pero hay 1 ocho en 10; $10 - 8 = 2$. Hay 0 cuatros en 2, pero hay 1 dos en 2; $2 - 2 = 0$. Hay 0 unos en 0. Concluimos que $106_{diez} = 1101010_{dos}$.

¿Cómo se expresaría 47_{diez} en el sistema binario? Hay 1 treinta y dos en 47; $47 - 32 = 15$. Hay 0 dieciséis en 15. Hay un ocho en 15; $15 - 8 = 7$. Hay 1 cuatro en 7; $7 - 4 = 3$. Hay 1 dos en 3; $3 - 2 = 1$. Hay 1 uno en 1. Por tanto, $47_{diez} = 101111_{dos}$. Pero este método es tedioso. Hay otro procedimiento que evita gran parte de esta labor.

Para convertir un número decimal a la forma binaria lo dividimos sucesivamente por 2. Por ejemplo, si dividimos 47_{diez} sucesivamente por 2, obtenemos:

$$\begin{array}{r} 23 \\ 2\overline{)47} \\ \underline{46} \\ 1 \end{array} \qquad \begin{array}{r} 11 \\ 2\overline{)23} \\ \underline{22} \\ 1 \end{array} \qquad \begin{array}{r} 5 \\ 2\overline{)11} \\ \underline{10} \\ 1 \end{array} \qquad \begin{array}{r} 2 \\ 2\overline{)5} \\ \underline{4} \\ 1 \end{array} \qquad \begin{array}{r} 1 \\ 2\overline{)2} \\ \underline{2} \\ 0 \end{array} \qquad \begin{array}{r} 0 \\ 2\overline{)1} \\ \underline{0} \\ 1 \end{array}$$

Los restos, en orden inverso, constituyen el equivalente binario de 47_{diez}. Leyendo desde el final hacia arriba, los restos son 1, 0, 1, 1, 1, 1. Por tanto, $47_{diez} = 101111_{dos}$. Esto coincide con el resultado obtenido usando el procedimiento más largo.

SUMA BINARIA

Supongamos que queremos sumar 11_{dos} y 110_{dos}. Esperamos que la suma represente el mismo número natural que si los números estuvieran expresados en base diez. Como en la base diez, la suma en base dos siempre es posible; la suma de dos números cualesquiera es única; la suma es conmutativa y asociativa, y cero representa el elemento neutro.

Recordando estas propiedades estructurales, es relativamente fácil desarrollar la suma en base dos. (En los ejemplos que siguen, suprimiremos a veces el subíndice "dos" para evitar confundir nuestra notación.) Comenzamos:

$$1 + 0 = 0 + 1 = 1$$

y

$$0 + 0 = 0$$

El único caso restante es $1 + 1$. Pero ya hemos visto que $1 + 1 = 10$ en el sistema binario. Por tanto, escribimos:

$$1 + 1 = 10$$

Podemos resumir nuestros casos de suma en forma de una matriz para fácil referencia.

+	0	1
0	0	1
1	1	10

Sumemos ahora $100 + 10$. Tenemos:

$$\begin{array}{r} 100 \\ +\ 10 \\ \hline 110 \end{array}$$

Ahora ensayemos $10 + 11$:

$$\begin{array}{r} 10 \\ +\ 11 \\ \hline 101 \end{array}$$

Discusión: $0 + 1 = 1$; por tanto, escribimos 1 en el lugar de los unos en la suma. Luego $1 + 1 = 10$. Escribimos 0 en el lugar de los dos en la suma, y 1 en el lugar de los cuatros.

Ahora solucionaremos dos ejemplos más difíciles, usando sólo el algoritmo de la suma convencional.

(a) $11010 + 10111 = \square$

(b) $110100 + 1111001 = \square$

Solución (a):

$$\begin{array}{r} 11010 \\ +\ 10111 \\ \hline 110001 \end{array}$$

Discusión: $0 + 1 = 1$; escribimos 1 en el lugar de los unos en la suma. $1 + 1 = 10$; escribimos 0 en el lugar de los dos y trasportamos 1 cuatro. $1 + 0 + 1 = 10$; escribimos 0 en el lugar de los cuatros y trasportamos 1 ocho. $1 + 1 = 10$; escribimos 0 en el lugar de los ochos y trasportamos 1 dieciséis. $1 + 1 + 1 = 11$; escribimos 1 en el lugar de los dieciséis y 1 en el lugar de los treinta y dos.

Solución (b):

$$\begin{array}{r} 110100 \\ +\ 1111001 \\ \hline 10101101 \end{array}$$

Discusión: $0 + 1 = 1$; escribimos 1 en la suma. $0 + 0 = 0$; escribimos 0 en la suma. $1 + 0 = 1$; escribimos 1 en la suma. $0 + 1 = 1$; escribimos 1 en la suma. $1 + 1 = 10$; escribimos 0 en la suma y trasportamos 1. $1 + 1 + 1 = 11$; escribimos 1 en la suma y trasportamos 1. $1 + 1 = 10$; escribimos 0 y luego 1 en la suma.

Aquí hay dos ejemplos más para usted. Vea si sus sumas concuerdan con éstas.

$$\begin{array}{r} 110101 \\ +\ 1100011 \\ \hline 10011000 \end{array} \qquad \begin{array}{r} 1101 \\ +\ 1001 \\ \hline 10110 \end{array}$$

RESTA BINARIA

La resta en el sistema binario se efectúa esencialmente lo mismo que en el sistema decimal.

Lo que se necesita con la notación binaria es alguna flexibilidad adicional.

De 110 restamos 101.

$$\begin{array}{r} 110 \\ -\ 101 \\ \hline 1 \end{array}$$

Discusión: No podemos restar 1 uno de 0 unos, por tanto tomamos una unidad de la siguiente posición (dos) y lo pensamos como "10" unos. "10" unos menos 1 uno deja 1 uno; escribimos 1 en el lugar de los unos en la diferencia. 0 dos menos 0 dos dejan 0 dos; 1 cuatro menos 1 cuatro deja 0 cuatros.

He aquí otro ejemplo:

$$1001 - 11 = \square$$

$$\begin{array}{r} 1001 \\ -\ 11 \\ \hline 110 \end{array}$$

Discusión: 1 uno menos 1 uno da 0 unos; escribimos 0 en el lugar de los unos en la diferencia. No podemos restar 1 dos de 0 dos. La posición a la izquierda de los dos es la de los cuatros, pero tampoco hay cuatros. Por tanto, vamos al lugar de los ochos. Cambiamos 1 ocho en "10" cuatros. Tomamos 1 unidad de los "10" cuatros, lo que nos deja 1 cuatro, y los cambiamos en "10" dos. Ahora 1 dos de "10" dos nos deja 1 dos; escribimos 1 en el lugar de los dos en la resta. Tenemos aún un cuatro restante, que colocamos escribiendo 1 en el lugar de los cuatros en la resta.

Realicemos un ejemplo más difícil:

$$10110 - 111 = \square$$

$$\begin{array}{r} 10110 \\ -\ 111 \\ \hline 1111 \end{array}$$

Discusión: 1 es mayor que 0. Cámbiese 1 dos en "10" unos. 10 menos 1 es 1. Cambiamos 1 cuatro en "10" dos. 10 menos 1 es 1. Cambiamos 1 dieciséis en "10" ochos. Tomamos 1 de esos ochos, que deja 1 ocho y lo cambiamos por "10" cuatros. Ahora completamos la resta. 10 menos 1 es 1. 1 menos 0 es 1. La diferencia es 1111.

Aquí hay dos ejemplos más:

$$\begin{array}{r} 10011100 \\ -\ 110101 \\ \hline 1100111 \end{array} \qquad \begin{array}{r} 10101101 \\ -\ 1111001 \\ \hline 110100 \end{array}$$

MULTIPLICACIÓN BINARIA

Supongamos que queremos multiplicar 10 por 11. Consideramos que el producto representa el mismo número que si la multiplicación se realizara en base diez. Como en base diez, la multiplicación en base dos siempre es posible, el producto de dos números cualesquiera es único, la multiplicación es conmutativa y asociativa y 1_{dos} representa el elemento neutro. También consideramos que las leyes distributivas se mantienen, o sea, la multiplicación es distributiva respecto de la suma.

Recordando estas consideraciones, desarrollemos la multiplicación en el sistema binario. Tenemos:

$$1 \times 0 = 0 \times 1 = 0$$

y

$$1 \times 1 = 1$$

Nuestra matriz de multiplicación sería entonces bastante simple.

×	0	1
0	0	0
1	0	1

Multipliquemos: 11 por 10

```
   10
 × 11
   10
  10
  110
```

Discusión: $1 \times 0 = 0$; escribimos 0 en el producto. $1 \times 1 = 1$; escribimos 1 a la izquierda del 0 en el primer producto parcial. $1 \times 0 = 0$; escribimos 0 bajo el 1, en el segundo producto parcial. $1 \times 1 = 1$; escribimos 1 a la izquierda del 0 recién escrito. Ahora sumamos.

Ahora estamos listos para abordar un ejemplo de multiplicación más difícil.

$$101 \times 1101 = \square$$

Solución:

```
    1101
  ×  101
    1101
   0000
  1101
 1000001
```

Discusión: Para el primer producto parcial: $1 \times 1 = 1$, $1 \times 0 = 0$, $1 \times 1 = 1$, y $1 \times 1 = 1$. Todas las cifras del segundo producto parcial son 0. Las cifras en el tercer producto parcial son iguales a las del primero. La suma es complicada, pero no demasiado difícil.

Aquí hay unos pocos ejemplos más para usted. Vea si sus productos concuerdan con éstos.

```
     10111          11100
   ×  1101        ×   111
     10111          11100
   10111           11100
  10111           11100
  100101011       11000100
```

DIVISIÓN BINARIA

Para poder dividir con números binarios uno debe tener práctica con la multiplicación y la resta binaria. Los primeros pocos ejemplos sobre los que trabajaremos se realizarán por restas sucesivas. Luego recurriremos a un algoritmo más convencional.

Dividamos 11011_{dos} por 11_{dos}.

```
       1001
 11) 11011
     11
      0011
        11
         0
```

Discusión: 11 está contenido en 11 justo 1 vez. $1 \times 11 = 11$. Restemos y resultará 0. Bajemos un 0. 11 está contenido 0 veces en 0; escribamos 0 en el cociente. Bajemos 1. 11 está contenido 0 veces en 1; escribamos 0 en el cociente. Bajemos 1. 11 está contenido 1 vez en 11. El resto es 0.

Veamos otro ejemplo:

```
        110
 101) 11110
      101
       101
       101
        00
```

Discusión: 101 está contenido en 111 sólo una vez. $1 \times 101 = 101$. Restemos y obtendremos 10. Bajemos 1. 101 también está contenido en 101 una vez. $1 \times 101 = 101$. Restemos y obtendremos 0. Bajemos 0. 101 está contenido en 0 cero veces.

Aquí hay dos ejemplos más para usted. Vea si sus cocientes concuerdan con éstos. Usted puede controlar su trabajo multiplicando el divisor por el cociente para obtener el dividendo.

```
              1110
 10011) 100001010
        10011
         11100
         10011
          10011
          10011

           1101
 1101) 10101001
       1101
       10000
        1101
         1101
         1101
```

DE FRACCIONES BINARIAS A DECIMALES

Convertir números fraccionarios en notación binaria a números decimales equivalentes no es particularmente difícil. El lugar de la derecha de la "coma binaria" tiene un valor de ½; el lugar siguiente tiene un valor de ¼; el siguiente, un valor de ⅛, etc. En cada lugar, el numerador es 1, y el denominador es la potencia de dos que depende del lugar; por ejemplo, el tercer lugar a la derecha de la coma binaria es $\frac{1}{2^3}$, ó ⅛. Consideremos ahora un ejemplo:

Ejemplo: $1{,}1011_{dos} = \underline{?}_{diez}$

Solución:

$$\begin{aligned} 1{,}1011 &= 1 \times 1 + 1 \times \tfrac{1}{2} + 0 \times \tfrac{1}{4} \\ &\quad + 1 \times \tfrac{1}{8} + 1 \times \tfrac{1}{16} \\ &= 1 + \tfrac{1}{2} + 0 + \tfrac{1}{8} + \tfrac{1}{16} \\ &= 1\tfrac{11}{16} \text{ ó } 1{,}6875_{diez} \end{aligned}$$

DE DECIMALES A LA FORMA BINARIA

Cualquier fracción decimal puede convertirse en su equivalente binario como sigue: Primero escribamos la fracción; luego la multiplicamos por 2 usando los casos de multiplicación en base diez. Por ejemplo, si la fracción decimal es 0,721, escribimos:

```
0,721
1,442
```

Luego quitamos el entero a la izquierda de la coma decimal y lo registramos, digamos a la derecha:

```
0,721
1,442   1
0,442
```

Repitiendo, tenemos:

```
0,884   0
0,884
1,768   1
0,768
1,536   1
0,536
1,072   1
0,072
```

El procedimiento se repite hasta que la fracción decimal se reduce a cero, o hasta que se obtenga el número deseado de lugares en la fracción binaria. El equivalente binario se obtiene leyendo los dígitos registrados del principio al final.

Así $0{,}721_{diez} = 0{,}10111_{dos}$.

DE NUMERACIÓN BINARIA A OCTAL

Un número expresado en forma binaria se expresa fácilmente como un equivalente en base ocho. Para un número natural, la técnica es simple. Como $2^3 = 8$, comenzamos con la menor unidad y agrupamos los "dígitos" en grupos de tres "dígitos" cada uno.

Ejemplos: Convertir 10110111_{dos} al número equivalente en base ocho.

Solución:

```
(10)    (110)   (111)dos = 267ocho
 ↑        ↑       ↑
2ocho   6ocho   7ocho
```

Ejemplo: Convertir, 1101101_{dos} a base ocho.

Solución:

$,(110)\ (110)\ (100)_{dos} = ,664_{ocho}$

RESUMEN

El sistema binario de numeración ofrece una interesante perspectiva en el mundo de las matemáticas. La suma y la multiplicación, a primera vista, parecen ser operaciones muy simples en el sistema binario, porque hay tan pocos

casos que aprender. Un estudio más cuidadoso revela que cualquier número de magnitud considerable requiere una sucesión casi interminable de ceros y unos para su representación. Las operaciones se vuelven muy rápidamente engorrosas en estos números largos y dificultosos. Pero las computadoras, a causa de su tremenda velocidad, pueden utilizar la notación binaria con gran ventaja. La mayoría de las computadoras digitales de diseños más antiguos emplearon notación decimal, pero pronto se hizo evidente que la notación binaria tenía muchas ventajas sobre otros sistemas de numeración para el diseño de circuitos de computación.

Hay muchos fenómenos interesantes que conducen por sí mismos a una explicación binaria. Si un asunto puede resolverse por "sí" o "no"; si un circuito está "cerrado" o "abierto"; si una lámpara está "encendida" o "apagada"; si una elección incluye "masculino" o "femenino"; si el resultado es "binario impar" o "binario par", los matemáticos aplican a este tipo de situaciones la notación binaria y la encuentran particularmente conveniente.

Los programas de "matemáticas modernas" que se enseñan comúnmente en algunas escuelas, tanto elementales como secundarias, tratan el tema de la numeración binaria con bastante detalle.

NIM: UN JUEGO INTERESANTE BASADO EN NÚMEROS BINARIOS

Un juego interesante para dos personas, denominado Nim, se basa en la numeración binaria. En este juego, los jugadores mueven por turno fichas de tres montones situados delante de ellos. Un jugador puede mover tantas fichas como quiera de cualquier montón en una sola jugada. En la siguiente jugada puede tomar fichas del mismo o de cualquier otro montón, según desee. El jugador que toma la última ficha que quedaba de los tres montones es el perdedor.

Si juegan un experto y un novato, el segundo difícilmente ganará. El secreto del juego, como un jugador experimentado sabrá, es seleccionar el número apropiado de fichas del montón correcto de tal modo que su disposición sea "binaria par". Esto fuerza la del otro a ser "binaria impar".

He aquí lo que significa "binaria par" y "binaria impar". Suponga usted que los montones de fichas designados como A, B y C contienen 9, 11 y 15 fichas respectivamente. Represente 9, 11 y 15 en forma binaria: $9 = 1001_{dos}$, $11 = 1011_{dos}$ y $15 = 1111_{dos}$.

A →	1	0	0	1
B →	1	0	1	1
C →	1	1	1	1

Una vez hecho esto, centre su atención en las columnas de números en lugar de la original representación en filas. Si la suma de cada columna es 0 ó 2, entonces su posición es "binaria par". Si no, su posición es "binaria impar". Cuando le llegue el turno de hacer una jugada, ojalá esté en posición "impar", porque si es par, cualquier movimiento que haga lo forzará a una posición y lo puede hacer perder... si su oponente conoce el juego.

Supongamos que usted se enfrenta con la situación planteada en la ilustración: 9, 11 y 15 fichas. Su posición es claramente "impar"; por tanto, debe realizar un movimiento que trasforma su posición en "par". Con un poco de práctica, verá que debería mover 13 de las fichas del montón C. Si lo hace, su posición aparecerá así:

A →	1	0	0	1
B →	1	0	1	1
C →			1	0

Claramente, la suma de todas las columnas es 2 ó 0. 0, en otras palabras, es binaria par.

Una vez que alcanza una posición segura (par), no interesa lo que haga su oponente: usted ganará, con tal que realice un movimiento "par" cada vez.

PROCESAMIENTO DE DATOS

¿Cuáles son las marcas de sopa más lucrativas? ¿Cuáles son las que hacen perder dinero? ¿Cómo determina el gerente de un supermercado las respuestas a preguntas como éstas? Primero reunirá *datos* –hechos y cifras– sobre las ventas de sopa en su negocio: hallará cuántas latas de cada marca se venden en cierto lapso; cuánto paga él por lata; cuánto paga el cliente por esa misma lata, etc. Esto se denomina recolectar, o presentar, datos, lo que debe hacer cuidadosamente: deben verificarse para tener certeza de que son precisos. Luego, el gerente dispondrá los datos en forma comprensible y definitiva, tal como con una gráfica o un diagrama: entonces, los datos –denominados *entrada*– estarán bajo tal forma que cualquiera podrá utilizarlos, o manejarlos.

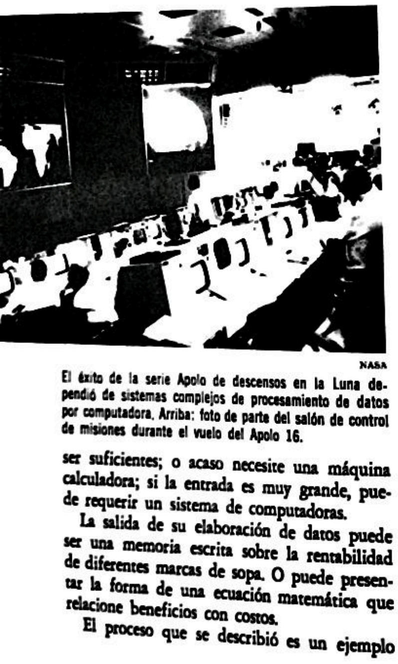
De esta entrada el gerente desea extraer información significativa sobre la rentabilidad de las diferentes marcas de sopa: esta información es la *salida* de su entrada de datos elaborados. El proceso *origen-entrada-elaboración-salida* es la piedra angular de todo el procesamiento de datos.

El gerente puede manejar la entrada de datos de varias maneras: papel y lápiz pueden

El procesamiento de datos a máquina es vital en la vida moderna. Tiene muchas aplicaciones, que van (hilera superior, izquierda a derecha) desde el análisis por computadora del corazón de un paciente, a la veloz confirmación de reservas de avión, y a la computación familiar de la caja registradora. Las computadoras procesan los registros de bancos (abajo, izquierda) y dan rápidamente todo tipo de datos comerciales (abajo, derecha).

UPI, IBM, NCR; Raymond Juschkus, Chase Manhattan Bank; NCR

El registro de los datos es esencial para el procesamiento de éstos. Se los puede almacenar en tarjetas perforadas (arriba y al centro), en cinta magnética (abajo, a la izquierda), o en discos magnéticos (abajo, a la derecha).



El éxito de la serie Apolo de descensos en la Luna dependió de sistemas complejos de procesamiento de datos por computadora. Arriba: foto de parte del salón de control de misiones durante el vuelo del Apolo 16.

ser suficientes; o acaso necesite una máquina calculadora; si la entrada es muy grande, puede requerir un sistema de computadoras.

La salida de su elaboración de datos puede ser una memoria escrita sobre la rentabilidad de diferentes marcas de sopa. O puede presentar la forma de una ecuación matemática que relacione beneficios con costos.

El proceso que se describió es un ejemplo de procesamiento de datos: el gerente (1) *originó* los datos; (2) los dispuso como *entrada*; (3) los elaboró y (4) produjo una *salida* de información ordenada.

Sin darnos cuenta, empleamos procesamiento de datos en nuestra vida diaria: por ejemplo, cada vez que vemos, oímos, olemos, paladeamos o sentimos algo, van datos al cerebro en forma de impulsos nerviosos; el cerebro los elabora para obtener la información que necesitamos para pensar o comportarnos de manera determinada. Así, la información y las reacciones que suscita pueden considerarse como la salida de la actividad cerebral.

En el sentido más amplio, un dato es cualquier cosa que produce información, la que significa datos que han sido depurados, resumidos y dispuestos en forma lógica.

Hay muchos modos de comunicar, o enviar datos; la palabra y la escritura son, quizá, los más comunes; otros son la taquigrafía, impulsos eléctricos y patrones magnéticos.

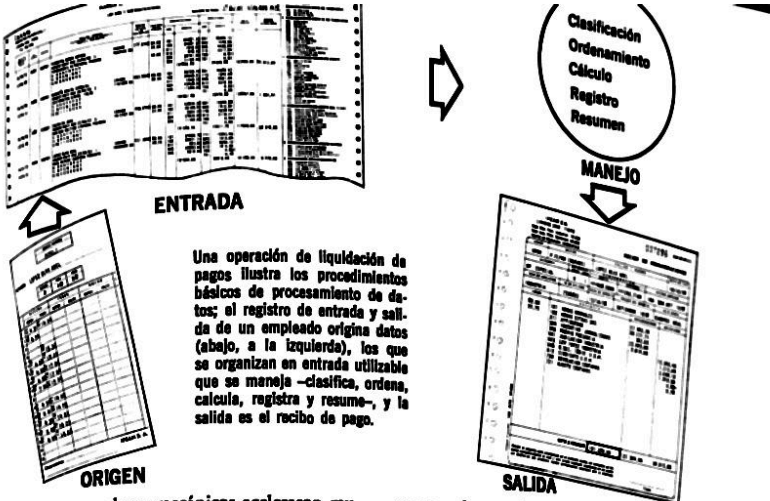
PROCESAMIENTO DE DATOS CON MÁQUINAS

Durante muchos siglos, la gente ha empleado lenguaje oral y escrito, y números, para procesar datos; la mayoría surge de actividades humanas como acontecimientos históricos, transacciones gubernamentales y financieras, vida social y diversos asuntos personales.

Hasta hace casi un siglo, la palabra y la escritura bastaban para procesar tales datos, pero esto resultó muy lento a medida que crecía la población y las instituciones se volvían más grandes y complejas. Entonces, el hombre buscó ayuda en las máquinas.

Al igual que los seres humanos, muchas máquinas –de escribir, calculadoras electrónicas, registradoras y ciertas partes de las computadoras– recurren a letras, palabras y números ordinarios. Pero otros dispositivos usan un "lenguaje" propio: orificios codificados perforados en tarjetas y cinta de papel; pulsos eléctricos y electrónicos; patrones magnéticos en cintas y discos, etc.

Desde cerca de 1870, las máquinas de procesamiento de datos comenzaron a adoptarse en gran escala. Estos primeros dispositivos eran estrictamente mecánicos; no empleaban electricidad ni electrónica. Muchos de ellos –tales como las máquinas registradoras y sumadoras– todavía son muy útiles hoy.



Una operación de liquidación de pagos ilustra los procedimientos básicos de procesamiento de datos; el registro de entrada y salida de un empleado origina datos (abajo, a la izquierda), los que se organizan en entrada utilizable que se maneja –clasifica, ordena, calcula, registra y resume–, y la salida es el recibo de pago.

Estos procesadores mecánicos aceleraron mucho el manejo de datos. Operaciones que insumían semanas y meses podían realizarse con mayor precisión en pocos días u horas.

Pero las instituciones gubernamentales y financieras crecieron y se complicaron aún más; se necesitó todavía mayor velocidad en el procesamiento de datos y, para satisfacer este requerimiento, se recurrió a la electricidad, la que hizo que las máquinas de escribir, de calcular y muchos tipos nuevos de dispositivos para las finanzas fuesen más rápidos y exactos. Los cálculos y otras operaciones con datos se redujeron de días a horas e, inclusive, a minutos. Hasta 1940, aproximadamente, estos llamados *procesadores electromecánicos* llenaron las necesidades del gobierno y de la industria; aún se emplean muchos de ellos.

Pero, desde 1940, la *computadora electrónica* revolucionó el procesamiento de datos. La primera aplicación seria de las computadoras fue la resolución de problemas militares complejos en la Segunda Guerra Mundial. Además se usaron computadoras pequeñas para apuntar cañones. Desde entonces, se difundió su uso en comercio, industria, gobierno, educación, ingeniería y otros campos.

Prácticamente no existe parte de nuestra vida que no haya sido afectada por las computadoras: ellas y dispositivos similares atienden, con frecuencia, nuestras facturas, cheques, préstamos, créditos y otras transacciones. Se las usa como ayuda para la enseñanza. Se puede determinar las necesidades económicas presentes y futuras de una nación por medio de computadoras, cuyas aplicaciones procesan y producen montañas de datos e información.

También han reducido enormemente el tiempo requerido para cálculos: cientos de miles de ellos pueden efectuarse en un segundo. Asimismo, procesamientos de datos que nunca antes se intentaron –a causa de su terrible dificultad– pueden ahora manejarse rápidamente gracias a ellas.

A diferencia de los procesadores de datos mecánicos y electromecánicos, el corazón de una computadora carece de partes móviles; empero, la entrada y salida de datos suele manejarse mecánica y electromecánicamente. Pero incluso estas fases se han acelerado recientemente.

PROCESAMIENTO DE UNA LIQUIDACIÓN DE PAGOS

Supongamos que usted tiene un empleo en alguna firma comercial o industrial: trabaja cierto número de horas por día, tantos días por semana y tantas semanas por año. Tiene cierto número de días libres por año: fines de semana, feriados reconocidos y vacaciones pagas. Además, quizá alguna vez no pueda trabajar todo un día, o tal vez esté enfermo.

La firma debe llevar un registro exacto del tiempo que usted trabaja y del que no lo hace: debe llevarlo no sólo para estimar su desempeño laboral, sino para calcular su salario. Usted recibe un jornal definido, basado en un índice horario de compensación. De éste, la compañía deducirá ciertos gastos del em-

pleado: impuestos, jubilación, seguro y beneficios sociales. Semanal, quincenal, o mensualmente, usted recibe un cheque de pago.

Estas *operaciones de liquidación de pagos* pueden ser complejas y difíciles, especialmente en compañías con cientos, o miles, de personas. En la descripción que sigue de una típica operación de liquidación de pagos suponemos que no se usa computadora: o sea, que el trabajo se hace a mano, o con máquinas de oficina electromecánicas comunes.

Al llegar a su trabajo, el empleado firma una hoja de entrada, indicando la hora en que se presenta, o perfora una tarjeta en un reloj de control. Así se registra la hora exacta de su llegada. Después del trabajo cotidiano, firma su hora de salida, o vuelve a perforar en el reloj la hora en que se retira.

De esta manera, ha *originado* datos sobre su asistencia diaria al trabajo. Al fin de la semana –u otro período conveniente– se comprueba la precisión de su registro de entrada y salida y del de otros empleados y se copia. Quizá se le dé entrada en alguna otra forma de registro; en otras palabras, los registros de tiempo se convierten a una forma comprensible de entrada de datos.

La entrada se *elabora* para obtener la salida; en este caso, una de las salidas más importantes es su cheque de pago. Como en muchas otras situaciones, la elaboración de la entrada de datos se lleva a cabo aquí en cuatro etapas básicas: (1) clasificación, (2) ordenamiento, (3) cálculo y registro y (4) resumen.

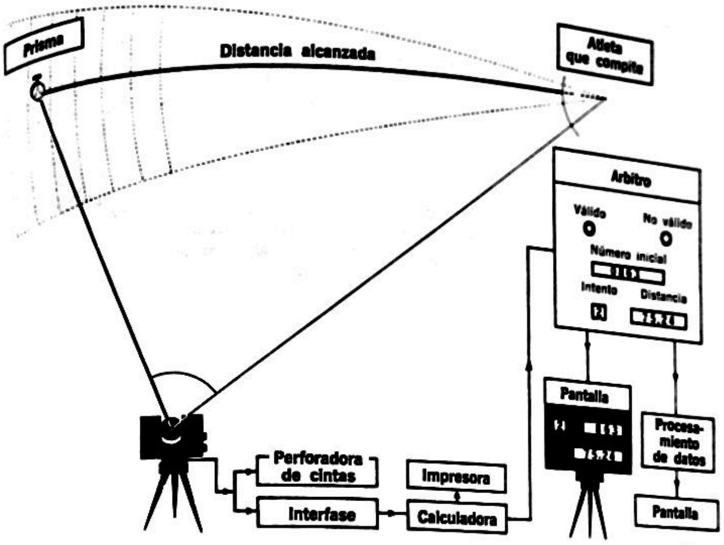
Veamos ahora cada etapa. Cuando los datos de entrada –registros de tiempo de los empleados– se *clasifican*, se los divide en diferentes grupos. Por ejemplo, los registros mencionados suelen agruparse según los departamentos de la compañía. Cada empleado de un departamento –como contaduría, producción o ventas– tiene normalmente un número.

Adviértase aquí un hecho importante: los datos que contienen nombre, número, departamento, emolumentos y deducciones de un empleado existen antes de que ingresen los datos actuales sobre su tiempo de trabajo. Estos datos anteriores –que, generalmente, se consideran permanentes– se comparan, o combinan, con los nuevos (tiempo de trabajo). El resultado es la nueva salida (cheque de pago del día). La reelaboración de los registros antiguos y recientes es uno de los aspectos más importantes del procesamiento de datos.

En las Olimpiadas se emplea el procesamiento electrónico de datos. Segundos después de que un atleta arroja la jabalina, la distancia alcanzada se mide con ayuda de un haz de luz, y el resultado aparece en un tablero. Un reflector manual de haces de luz se muestra en la parte superior de la foto.

Después de la clasificación, la entrada se *ordena*. Dentro de cada grupo –en este caso, un departamento de la compañía–, los datos se colocan en cierto orden o secuencia. Los registros de tiempo de los empleados pueden disponerse en orden alfabético (por nombre), o numérico (según el número de cada uno).

Una vez clasificados y ordenados, los empleados pagadores *calculan* y *registran* jornales, horas trabajadas, tiempo no trabajado, y las deducciones para impuestos, jubilación, seguro y beneficios sociales. Otra fase importante del manejo es la adición, o *resumen*, de los datos que se procesan. Gran parte de los atinentes a liquidación de sueldos son muy



Medición de las distancias de las pruebas deportivas en las Olimpiadas. Después de cada lanzamiento, un prisma marcó el punto de impacto. Un instrumento de medición (abajo, izquierda), colocado a distancia conocida del atleta, emitió entonces un haz de luz que, al reflejarse, dio la distancia entre el prisma y el dispositivo medidor. Luego, los instrumentos computaron el tercer lado del triángulo: la distancia alcanzada.

importantes para los gerentes de una firma, quienes usan las cifras para calcular costos de mano de obra de la compañía en un período dado. Pueden así tomar importantes decisiones sobre la política que debe seguir la compañía.

Después de la producción de datos, su entrada y elaboración, tenemos una salida; en nuestro ejemplo, esto incluye –entre otras cosas– su cheque de pago, el que, por lo usual, viene en dos partes adjuntas: una indica su paga neta (paga después de todas las deducciones), y es la que usted endosa y luego cobra en el banco, o deposita en su propia cuenta. La otra parte del cheque indica su paga bruta, y puede enumerar las deducciones –seguro, jubilación, impuestos y beneficios sociales– que se descontaron, de lo que resultó su paga neta.

PROCESAMIENTO DE DATOS CON COMPUTADORA

Básicamente, el procesamiento de datos con computadora es casi lo mismo que el que se realiza a mano, o por métodos electromecánicos; la diferencia principal estriba en que la computadora maneja toda la tarea simultáneamente, en una operación continua a alta velocidad. Ciertos pasos, tales como el ordenamiento, pueden dejarse completamente de lado. Muy poco es lo que hace el ser humano.

El procesamiento de datos por computadora tiene lugar en tres fases: *entrada*, la operación de *procesamiento* en sí y la *salida*. La entrada de datos se hace en forma tal que la máquina pueda "leerla" para procesarla. Los datos pueden registrarse en tarjetas o cinta perforadas, en cinta magnética, o en tambores o discos magnéticos; o pueden suministrarse a la máquina como números, palabras, letras u otros símbolos impresos o mecanografiados.

Luego, la computadora procesa la entrada según un programa que se le introdujo previamente. En el caso de una operación de liquidación de pagos, la computadora lee todos los datos necesarios: nombre del empleado, su

número, departamento, índice de pagos, nú-
mero de horas trabajadas, tiempo no trabajado,
deducciones, etc. Entonces calcula los sueldos
bruto y neto del empleado. Estas cifras se des-
cargan de la memoria de la máquina –o depó-
sitos de almacenamiento de datos– y se im-
primen en un cheque.

En otras computadoras, la salida no se im-
prime; en cambio, se registra como código
sobre tarjetas perforadas, cinta, cinta magné-
tica, o tambores o discos magnéticos.

ARCHIVO DE LOS DATOS

Un archivo de datos es un conjunto de re-
gistros. Existen archivos sobre casi cualquier
tópico: nacimientos, muertes, crímenes e in-
vestigación científica.

Un tipo familiar de archivo es una colec-
ción de documentos escritos e impresos –pa-
peles, tarjetas, folletos, recortes y, a veces, li-
bros– que, usualmente, se mantienen en un
orden definido y se almacenan en gabinetes,
gavetas o estantes.

Otras clases de archivos pueden resultar me-
nos familiares. Los datos pueden fotografiarse
en microfilmes, muy pequeños para leerse a
simple vista, pero que ocupan muy poco

"¿Dónde siente dolor?" Un paciente responde usando una "lapicera de luz" electrónica. Una computadora procesa luego los datos médicos del paciente e imprime un resumen para que lo estudie el facultativo.

IBM

espacio de almacenamiento. Otros archivos
pueden contener datos codificados como ori-
ficios perforados en tarjetas o cinta de papel.
Otros, aún, pueden registrarse como patrones
en tambores, discos o cinta magnéticos.

La información –o salida terminada de da-
tos procesados– a menudo se almacena en
archivos, como se hace con aquéllos. Pero la
información proveniente de un sistema de pro-
cesamiento de datos puede a su vez servir co-
mo dato para otro sistema similar; así, a me-
nudo se cambian los archivos de una clase a
otra. Por ejemplo, documentos impresos de un
archivo pueden codificarse en tarjetas perfo-
radas o cinta magnética y usarse luego como
entrada para una computadora.

Cualquiera sea su forma, un archivo tanto
almacena registros como los pone a disposi-
ción de quienquiera que los precise. Por ejem-
plo: una computadora mantiene registros en
su memoria, o depósitos de almacenamiento
de datos; usa estos datos más permanentes pa-
ra procesar los más recientes que ingresan: a
este depósito de datos se lo llama *almacena-
miento interno*. Al de datos fuera de la com-
putadora, *almacenamiento externo;* este último
a menudo existe en forma de cinta magnética,
tarjetas perforadas y documentos.

Hay dos maneras de clasificar un archivo:

1. Un *archivo funcional* se establece con
un propósito definido; los datos se clasifican
de acuerdo con su función o uso.

2. Un *archivo dispuesto físicamente* no se
establece para una función específica; sus re-
gistros se colocan en cualquier orden, o ubica-
ción, convenientes.

Un archivo funcional puede ser *maestro* o
de transacción: el primero sólo contiene
registros permanentes, relacionados, que de
cuando en cuando necesitan actualizarse. En
contraste, uno de transacción sólo contiene
registros recientes –o transacciones– que re-
presentan los cambios que deben realizarse
para poner al día un archivo maestro. El ar-
chivo de transacción es temporario.

Un archivo clasificado según la distribución
física mantiene sus registros en una de dos
formas: seriada, o de acceso al azar.

Seriado. En un archivo seriado, los regis-
tros se almacenan en orden de serie. Una má-
quina examinará todos los registros previos
antes de alcanzar el que se desea.

El medio más importante para el almace-
namiento de datos seriados es la cinta mag-
nética: supongamos que tenemos 10 registros
"escritos" en cinta en la sucesión numérica
adecuada: 01, 02, 03, 04, 05, 06, 07, 08, 09
y 10; si, por ejemplo, se necesitara el registro
06 para el procesamiento, la máquina debe
examinar los registros desde el 01 hasta el 05
antes de obtener el 06.

El almacenamiento en archivo seriado es
ideal cuando se necesita llevar a cabo la reac-
tualización regular de todos los registros, co-
mo en el caso de las cuentas de clientes. No
se pierde tiempo haciendo que la computa-
dora tenga que investigar los registros indi-
viduales de una cinta; simplemente, las cuen-
tas se han registrado una a continuación de
la otra, y la computadora las recorre y pone
al día rápidamente, en períodos regulares.

Acceso al azar. En estos archivos, los re-
gistros se almacenan generalmente en discos
magnéticos, muy similares a los fonográficos
de larga duración, excepto que cada banda del
disco magnético es una circunferencia, y todas
se disponen en un esquema de circunferencias
cada vez menores y concéntricas.

Los registros de una banda particular –o
sección de banda– tienen un número de có-
digo determinado: al suministrarlo una per-
sona que busca un registro, la máquina puede
ir directamente a la banda, o sección, que se
pide. En vez de recorrer todos los registros
archivados, sólo "leerá" los pocos con el mismo
número de código, para hallar el deseado.

PROCESAMIENTO DE LOS ARCHIVOS DE DATOS

Puesto que de muchos archivos se extrae la
información constantemente, deben mantener-
se actualizados. A medida que se realizan nue-
vas transacciones, los datos deben hacerse in-
gresar en el archivo: según sea éste, la entrada
se hace a máquina o por una persona. Si los
registros están almacenados en cintas, discos o
tambores magnéticos, o en cintas o tarjetas
perforadas, los datos deben ser leídos por una
máquina tal como una computadora.

Los archivos seriados se actualizan mediante
una operación denominada procesamiento en
bloques o por tandas. Los archivos de acceso al
azar se ponen al día mediante el procesa-
miento en línea.

Litton

No todas las máquinas de procesamiento de datos son voluminosas: esta calculadora electrónica y muchas otras –algunas inclusive más pequeñas– pueden llevarse en el bolsillo. Tales máquinas resuelven problemas aritméticos, exhibiendo la respuesta en la ventanilla de la parte superior.

Procesamiento en bloques. Los registros al-
macenados en serie sobre cinta magnética, o
tarjetas perforadas, requieren procesamiento
periódico. Durante cierto período se reúnen
registros de transacciones; luego, una tanda de
tales registros se procesa de una vez.

El procesamiento en bloques es ideal para
liquidaciones de pagos, cuentas comerciales y
otras transacciones seriadas. Hay un ahorro
evidente de tiempo de computación si éstas
pueden manejarse de este modo.

Un archivo maestro también puede proce-
sarse en bloques: si está en cinta magnética,
puede colocarse en una computadora cuando
requiera actualización. Previamente al acciona-
miento de la computadora, la tanda de datos
nuevos se ordena en el mismo orden que el
archivo maestro. Cuando comienza la operación,
la máquina lee cada fracción de datos y sitúa
el registro correspondiente en la cinta maes-
tra, operación que continúa hasta que la cinta
está reactualizada: se produce, así, una nueva
cinta maestra corregida, que se usa para la si-
guiente reactualización.

El procesamiento en bloques presenta varias
ventajas; como se procesa simultáneamente un

gran volumen de registros, hay un ahorro de tiempo de computadora y de dinero. Asimismo, disminuyen el tiempo y el gasto de preparar estos registros para procesarlos. Hay un control más efectivo de los errores. Pueden obtenerse totales de las cifras antes de cada etapa y después de ella, asegurándose así que no se han omitido transacciones.

Pero el procesamiento en bloques presenta también desventajas. Es difícil procesar datos en momentos distintos a los programados para aquella actividad. La puesta al día inmediata de un registro puede no ser posible, o ser muy cara. Otra desventaja es que impone a la computadora un trabajo pesado que debe realizar en un breve lapso. Ocasionalmente pueden surgir problemas serios, tales como demoras o paralizaciones.

Procesamiento en línea. Se usa un tipo de procesamiento enteramente diferente para los archivos de datos de acceso al azar. El procesamiento en línea maneja las transacciones tales como llegan, sin considerar su orden; no se necesita ordenamiento o agrupación en bloques antes de accionar la computadora.

El procesamiento en línea efectivo requiere el uso de dispositivos rápidos, llamados de *acceso directo*; son discos, o tambores, magnéticos, en los que puede ubicarse directamente cada registro o la localización del almacenamiento de datos.

En el sistema lineal, un registro maestro se pone al día poco después de ocurrir una nueva transacción. Este sistema es ideal cuando aquéllas no vienen en sucesión; es especialmente útil cuando la situación requiere una actividad más frecuente que la que se efectúa en el procesamiento en bloques.

Otras ventajas de la operación en línea incluyen un suministro más rápido de información, pues las transacciones se procesan a medida que llegan, y no se sobrecarga la computadora con excesiva cantidad de trabajo de una vez, de modo que no se demora el manejo de otros datos. El método generalmente más rápido de procesamiento en línea hace la operación eficiente, por el tiempo ahorrado.

Pero el procesamiento en línea presenta ciertas desventajas. Se necesitan grandes y costosos dispositivos de almacenamiento de datos. Los costos de procesamiento también pueden ser elevados, ya que, por lo común, se procesan relativamente pocas transacciones por vez. Asimismo, es difícil rastrear el flujo de datos de entrada a salida a través de un sistema lineal. Si el sistema se descompone, ello acarreará demoras serias y errores costosos.

Un policía disca un mensaje codificado, el que se trasmite, a través de una radio común, a una terminal donde una computadora lo procesa.

PROCESAMIENTO DE RESERVAS EN LÍNEAS AÉREAS

El procesamiento en línea es útil en especial para reservar pasajes de avión. El tiempo real que insume la obtención de una respuesta a una solicitud es cuestión de segundos.

Por ejemplo, la American Airlines posee el SABER, sistema de Investigación Semiautomática del Medio Mercantil, que permite a los agentes de venta reservar pasajes casi instantáneamente por medio de un complejo central de computación.

Cuando un presunto pasajero llama pidiendo una o más reservas en un vuelo, el agente de ventas perfora información codificada en una terminal, que se conecta –junto con otras– a un intercambio terminal situado en la misma zona, por líneas de baja velocidad.

El intercambio retiene el mensaje del agente hasta que el complejo central de computación en Nueva York (centro de operaciones), está listo para tomarlo; cuando esto ocurre, el mensaje llega al complejo a través de líneas de alta velocidad. Entonces, la computadora comienza a buscar el dato archivado en sus discos magnéticos para responder a la solicitud de pasajes.

Cuando la respuesta está lista, se trasmite desde el complejo de computación al agente que envió el mensaje original, el que, a su vez, informa al presunto pasajero si hay asientos en el vuelo deseado y si están confirmadas las reservas. Toda la operación, del principio al fin, insume de 3 a 5 segundos.

El programa empleado para el funcionamiento del SABER consiste en más de 150.000 instrucciones. Entre otras cosas, atiende los siguientes puntos claves de información: nombre del pasajero; nombre de la persona que hace la reserva; número de teléfono en el que puede encontrarse al pasajero; nombre de la persona que recoge los pasajes; solicitudes de menú especial; solicitudes de silla de ruedas; espectáculos especiales contratados en el destino del vuelo; solicitudes de alquiler de auto, etcétera.

La complejidad de este sistema se evidencia porque el inventario de tal información puede actualizarse, a partir de hasta un año antes de la fecha actual de reserva.

En la actualidad, agentes en todo el mundo operan las terminales de SABER y otros sistemas similares, tales como el PANAMAC, de la Pan American. Procesan cientos de miles de llamadas telefónicas diarias atinentes a reservas, preguntas sobre asientos, ventas de pasajes y manejo y trasporte de carga.

PROYECTO DEL PROCESAMIENTO DE DATOS

Cualquier empresa, no importa su magnitud, procesa datos. Si la salida ha de tener significado, esto debe hacerse sistemáticamente.

Un sistema de datos cumple dos tareas principales: crear datos para archivos y mantener a éstos actualizados. Así, de tiempo en tiempo, se crea una nueva salida, tal como cheques en una operación de liquidación de pagos.

Para que una organización obtenga la salida deseada, debe establecerse un método o procedimiento definido, con computadoras o sin ellas. Éste es el trabajo de los ***analistas de sistemas***. Estos expertos trabajan en equipo y discuten los problemas implicados.

El proyecto y análisis de sistemas incluye las actividades siguientes:

1. Recolección de datos: mostrar los que se han de utilizar.
2. Análisis de los datos: mostrar cómo debe manejárselos.
3. Proyecto de sistemas: crear el método a seguir para obtener la salida necesaria.

El analista de sistemas es la persona clave en la concepción de éstos; dirige a los encargados del desarrollo de los procedimientos necesarios; debe mantener informados a los gerentes sobre el sistema de procesamiento de datos que sirve mejor sus intereses, y respecto al estado del sistema en todo momento.

Además, el analista explora nuevos métodos de diseñar sistemas más eficientes de procesamiento de datos y se mantiene informado de las aptitudes y costos de los diversos equipos que se venden en el mercado. El analista debe justificar la necesidad de informes actualizados, suprimir procedimientos innecesarios, y pasar el personal del trabajo rutinario y aburrido a tareas más creativas.

Para satisfacer las normas y necesidades tratadas antes, debe escogerse cuidadosamente el equipo de procesamiento de datos. El mejor sistema es el que produce la salida deseada a la velocidad requerida, y a los costos más bajos posibles en tiempo, mano de obra y dinero.

También son importantes la magnitud de la empresa, el volumen de datos a procesar, la complejidad de las operaciones y la precisión necesaria; antes de elegir un sistema dado deben sopesarse todas estas consideraciones.

El sistema elegido puede ser manual, podría ser ideal para empresas pequeñas que manejan cantidad limitada de datos sencillos.

O bien el sistema puede usar equipo mecánico o electromecánico, como procesadores de tarjetas, o cintas perforadas; calculadoras y máquinas de escribir convencionales.

Una organización grande que maneje toneladas de datos necesita un sistema de computación. La operación principal escogida puede ser el procesamiento en bloque o en línea.

En suma: los esfuerzos combinados de hombre y máquina posibilitan que las organizaciones manejen datos en forma eficaz. Esta clase de sistema beneficia también a los individuos. Los obreros de las grandes compañías reciben a tiempo sus cheques de pago, gracias a las complejas máquinas de procesamiento de datos. Los matemáticos y científicos procesan datos en computadoras, completando así cálculos complejos en minutos, en vez de meses. Los bancos utilizan sistemas de procesamiento de datos para determinar el interés de las cuentas de ahorro. En éstos, y muchos otros casos, la gente que debe manejar cifras grandes se ve liberada de una faena penosa y casi interminable.

COMPUTADORAS

"Buenos días, Alicia." Mientras un terminal de computadora imprime este saludo, Alicia, de 8 años, empieza su lección de matemáticas modernas. Al mismo tiempo, una de sus compañeras de clase comienza una lección de ortografía con ayuda de otra computadora.

A tres mil kilómetros de distancia, un criador de aves de Oregón usa una de estas máquinas para calcular cuánto le costará duplicar el volumen de su explotación.

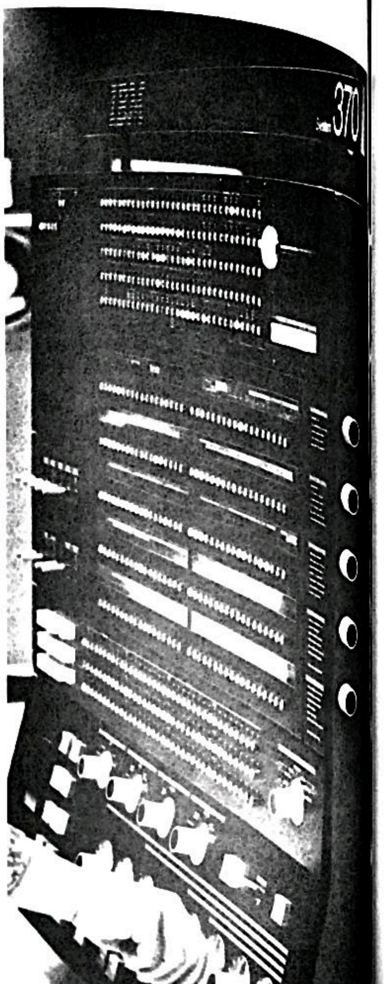
En la ciudad de Nueva York, un patrullero de la policía observa un vehículo sospechoso. El policía llama a la jefatura y trasmite a un operador de computadora el número de la patente. En segundos obtiene una respuesta: el auto no es buscado.

En la ciudad de México, un médico debe tratar a un turista norteamericano seriamente enfermo. El facultativo telefonea a un servicio de computación de historias clínicas de Nueva York. En unos segundos averigua que el paciente es alérgico a las sulfamidas, tiene grupo sanguíneo B^+ y es diabético.

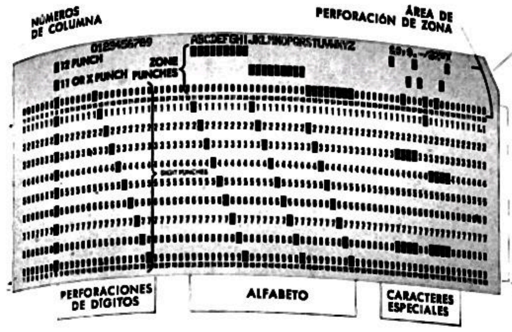
En Bélgica, un joven muere en un accidente automovilístico. Sus tejidos se clasifican y comparan por computadora con los tipos de tejido de 300 pacientes a la espera de un trasplante de riñón. Uno de los riñones del joven se envía a un hospital de Luxemburgo; el otro, a Inglaterra. El procedimiento completo, incluyendo las operaciones de trasplante, lleva menos de 12 horas.

Actualmente, las computadoras se usan para enseñanza de idiomas, construcción de aviones, rastreo de órdenes de fabricación, diseño de mapas y gráficas y aun búsqueda de cámaras ocultas en las antiguas pirámides de Egipto. En resumen, se han convertido en parte integrante de nuestra vida, así seamos científicos, hombres de negocios, estudiantes o amas de casa. Estas calculadoras electrónicas se han vuelto "cerebros gigantes" de los que dependen grandes sectores de la sociedad moderna.

La computadora, sin embargo, es básicamente simple y no se debe desconfiar de ella, ya sea por temor o desconocimiento. Pues simplemente acepta la entrada de algunos datos, los



La computadora sistema IBM/370 es capaz de desarrollar hasta 15 programas de computación simultáneamente.



Un medio usual de introducir la información en la computadora es la tarjeta perforada. Los informes aparecen como perforaciones en la tarjeta. Una unidad lee esta información y la trasfiere al almacenamiento de la computadora.

almacena, los maneja, selecciona entre varias líneas posibles de trabajo y luego exhibe la información en un dispositivo de salida.

Hay dos tipos de computadoras: digitales y analógicas. La ***computadora analógica*** trata directamente con cantidades mensurables como fuerzas, voltajes, presiones, temperaturas y otras variables continuas. Está proyectada para servir de analogía o semejanza física del problema por resolver. Se supone que las ecuaciones que describen las relaciones entre las variables de la computadora están estrechamente relacionadas con las variables del problema. Las trasformaciones en las variables de la computadora se interpretan como los resultados que producirían las trasformaciones en las variables del problema.

La ***computadora digital*** trabaja con números o símbolos codificados en forma numérica. El nombre de computadora digital viene de dígito, palabra que significa número de una cifra. Básicamente, las computadoras digitales cuentan cosas para obtener resultados. La calculadora de mesa es un dispositivo digital en contraposición con una regla de cálculo, que es un dispositivo analógico. Ambas pueden usarse para operaciones aritméticas tales como multiplicación o división.

Sin embargo, cuando la gente usa la palabra "computadora" se refiere generalmente al tipo digital. A pesar de que el tipo analógico es importante en navegación, gobierno de proyectiles, control de fuego de artillería y algunos procesos industriales, este artículo trata especialmente de la computadora digital.

PARTES DE UNA COMPUTADORA ELECTRÓNICA

Las computadoras usadas actualmente en nuestra sociedad son muchas y variadas. Cada una de ellas tiene ciertos componentes básicos que están siempre presentes, pero difieren en el número y variedad de unidades que se agregan para formar un sistema de computación. Conocer las partes básicas de funcionamiento da un punto de referencia para abordar los sistemas de computación más complejos. Tal vez, la parte más importante de una computadora es la ***unidad central de procesamiento.***

La unidad central de procesamiento (UCP) proporciona a la computadora aptitudes aritméticas, lógicas y de control. La unidad aritmética de la UCP se encarga de las operaciones fundamentales, a saber: suma, resta, multiplicación y división. Las operaciones lógicas son muy poco más que una capacidad para comparar dos números y determinar si son iguales; y si son distintos, cuál es el mayor.

La unidad de control del procesador central gobierna electrónicamente la operación de la computadora. Coordina las distintas unidades que la constituyen y determina en qué orden y por cuánto tiempo estas unidades deben activarse. La unidad de procesamiento lleva a cabo sus tareas ejecutando instrucciones de un programa. Con cada instrucción, la computadora avanza un paso muy pequeño en la realización de algún proceso. El conjunto de las instrucciones que componen las órdenes completas se denomina ***programa de computadora.***

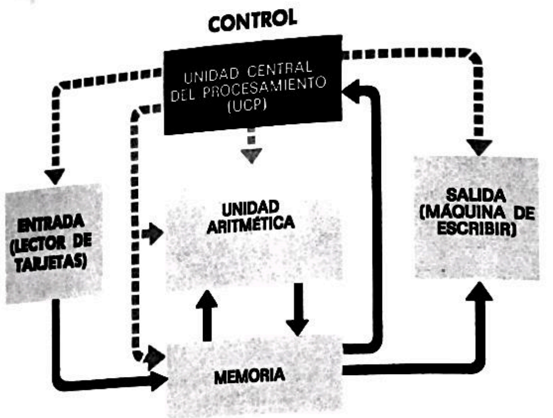


Figura 1. Las cinco partes básicas de una computadora. Las líneas llenas indican la dirección del flujo para los datos, valores y las instrucciones. Las líneas punteadas, el control que la UCP tiene sobre otras unidades.

Muy estrechamente asociada a la unidad central de procesamiento hay otra de almacenamiento. Es la unidad primaria de almacenamiento de la computadora y a menudo se denomina *memoria*. Esta unidad recibe datos, los mantiene indefinidamente sin borrarse o perderse, y proporciona la información requerida por la UCP para procesarlos. El almacenamiento primario y la UCP son, conjuntamente, el corazón del sistema de computación.

Las unidades de entrada y de salida son los mecanismos por cuyo intermedio los datos y otras informaciones llegan al depósito de la computadora y salen al exterior. Un medio común de entrada es la tarjeta perforada. Los datos, expresados por perforaciones en una tarjeta de cartulina, son leídos por la unidad lectora de tarjetas dentro del depósito de la computadora. Un dispositivo de salida muy importante es algún mecanismo impresor, que puede ser tan simple como una máquina de escribir, o tan complicado como una prensa de alta velocidad que imprimirá 2.000 líneas de 132 caracteres por minuto.

Finalmente, se ejerce el control humano de la computadora por medio de la consola. Luces, llaves y botones no sólo permiten al operador controlarla, sino también vigilar lo que ésta está haciendo en cualquier momento. Muy a menudo, la consola se proyecta junto con la UCP para que el procesador y la consola estén en una misma caja.

En la Figura 1 se muestra un diagrama simplificado de una computadora. Representa una pequeña, con una perforadora de tarjetas como entrada y una máquina de escribir como salida. Las líneas llenas indican la dirección del flujo de datos, valores e instrucciones. Por ejemplo, los valores de entrada se trasfieren de la lectora de tarjetas a la memoria; los de salida, de la memoria a la máquina de escribir. También hay trasferencia de valores entre la memoria y la unidad aritmética, en ambas direcciones. Las instrucciones se trasfieren de la memoria a la unidad de control. Las líneas cortadas indican la acción de controlar; por tanto, conectan la unidad de control con las unidades de entrada, salida y aritmética, así como con la memoria.

Cualquier computadora tiene estas partes básicas de funcionamiento. Los sistemas más grandes y complejos tienen mayor variedad de unidades que las mostradas en la Figura 1, y más de cada unidad.

ALMACENAMIENTO DE LA INFORMACIÓN

La computadora digital presenta los números en forma binaria. Un código binario popular es el Código Alternado Decimal Codificado Binario Ampliado, cuya abreviatura es CADCBA.

En este código, ocho dígitos binarios, u ocho *bits*, codifican un carácter. Cada grupo de ocho bits se denomina *byte*. Cada carácter que la computadora puede expresar se codifica en un único byte. Ocho bits proporcionarán 256 de estos códigos únicos, que pueden usarse para expresar 256 caracteres distintos en el almacenamiento de la computadora. Una lista seleccionada de caracteres CADCBA y sus códigos se muestra en la Tabla 1. La unidad de memoria almacena todos los datos en esta u otra forma binaria.

TABLA I

Caracteres codificados CADCBA seleccionados

espacio	01000000
.	01001011
+	01001110
–	01100000
=	01111110
A	11000001
B	1/000010
C	11000011
Z	11101001
D	11110000
1	11110001
2	1/110010
9	11111001

La unidad de memoria de una computadora debería tener cinco cualidades deseables: (1) ser confiable, en el sentido de que un carácter no se borre fácilmente; (2) ser grande, para poder almacenar muchos datos; (3) poder organizarse en lugares independientes, así los caracteres no se situarían unos encima de otros; (4) ser rápida para que el acceso a valores tome muy poco tiempo, y (5) ser barata.

No todas estas virtudes pueden encontrarse en un solo órgano de almacenamiento. Una forma de memoria muy popular se compone de *núcleos magnéticos*, que son confiables, rápidos y organizables en lugares independientes. Cuando las memorias de núcleo son grandes, también son caras; por tanto, el bajo costo se obtiene sólo con un tamaño limitado.

Un núcleo magnético se parece a una rosquilla de 0,5 milímetros de diámetro con un agujero de 0,4 milímetros (véase la Figura 2). Está hecho de un material ferroso que puede magnetizarse, ya sea en el sentido de las agujas del reloj o en el sentido contrario. Por ello, el núcleo puede almacenar información. Los dos "estados" magnetizados son únicos y discernibles y pueden representar los valores 0 y 1 de un dígito binario. Aunque un núcleo sólo puede almacenar el valor binario 0 ó 1, un conjunto de núcleos almacenará un gran número de bits.

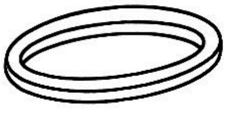
Una corriente eléctrica que circula por un cable genera un campo magnético alrededor de él (Figura 3*a*). Si la dirección de la corriente se invierte, también lo hace la dirección del campo magnético (Figura 3*b*). Si un núcleo magnético se inserta en un cable, el campo magnético generado por una corriente que circule por el cable magnetiza el núcleo en la misma dirección. Si una hilera completa de núcleos se somete al campo magnético generado por la corriente que circula por el cable, cada uno se magnetizará. Para su uso en la memoria de una computadora debe magnetizarse un solo núcleo en vez de la hilera completa. Esto se logra empleando las propiedades magnéticas del material que compone el núcleo.

Si la corriente eléctrica que circula por el cable es demasiado débil, el núcleo no se magnetiza en forma permanente. Sólo si la corriente es suficientemente elevada generará un campo magnético capaz de magnetizarla permanentemente, prescindiendo de su condición anterior. Una corriente que sea demasiado débil no lo afectará. Por esta razón, cada núcleo se ubica en la intersección de dos cables. Los cables se enhebran a través de los agujeros de los núcleos. La energía de la corriente que circula por los dos cables se regula con cuidado para que la de cada uno no alcance a magnetizar un núcleo. Sólo con dos pulsaciones eléctricas simultáneas la energía de la corriente excederá la requerida para magnetizar un núcleo permanentemente. La corriente que fluye a través de ambos cables magnetizará permanentemente sólo aquel núcleo que está ubicado donde éstos se cruzan (Figura 4).

La Figura 5 es una fotografía de un plano de núcleos de una computadora. Cada núcleo del plano está ubicado en una intersección dada de dos cables: uno tendido horizontalmente y otro verticalmente en el plano. Si ocho de estos planos se apilan unos sobre otros, la columna vertical de ocho núcleos en una misma intersección fila-columna de cada plano representa los ocho bits de un byte y por ende un carácter de datos codificados en binario.

Conjuntos de bits (núcleos) dispuestos en planos y pilas de planos forman capacidades de memoria de una computadora que van desde

FIGURA 2



unos 4.000 bytes hasta varios millones. Estos bytes tienen señas para ser identificados en forma única por el procesador central, de modo que los caracteres pueden almacenarse en la memoria independientemente y no dispuestos unos sobre otros. La seña es un número único asignado a cada grupo de ocho bits (una columna vertical de ocho núcleos).

CARACTERES BINARIOS

Internamente, los caracteres se representan como números binarios de ocho dígitos. Cada carácter tiene un valor numérico binario que lo distingue de cualquier otro. En los caracteres ilustrados en la Tabla 1, "espacio en blanco" tiene el menor valor numérico (01000000), y el dígito "nueve" tiene el mayor valor numérico (11111001). Estos valores no sólo identifican los caracteres individuales, sino que permiten las comparaciones de uno con otro. Por ejemplo, el mayor carácter especial es el "=", que significa (01111110). Numéricamente, es menor que la primera letra del abecedario, A (11000001). Z tiene mayor valor que cualquier otro carácter alfabético (11101001) y es menor que el menor de los dígitos decimales (11110000).

Comparar caracteres en la computadora es diferente a hacerlo en la lectura visual, donde la forma del carácter es esencial. En la computadora lo importante es el código interno binario. Dichos códigos se ordenan de modo que los valores numéricos binarios se sucedan en un orden que permita compararlos, y por tanto clasificarlos en una secuencia inteligible. Los valores que son nombres de personas, por ejemplo, se clasifican comparando los códigos de bytes en valor creciente para ordenarlos de A a Z, o sea, de 11000001 a 11101001.

En algunas aplicaciones y usos de la computadora es necesario conocer el contenido exacto de la memoria. Por contenido exacto se entiende la posición relativa de los bits individuales del valor de los datos. Una impresión del contenido de la memoria se denomina *descarga*. Una descarga de la memoria en forma binaria es muy difícil de interpretar. Aun es difícil comparar un byte de ocho bits con otro para ver qué representa cada uno. Es imposible buscar entre cientos y quizá miles de tales valores binarios, aunque a veces sea necesario. Por tanto, las computadoras organizadas en bytes son generalmente capaces de representarlos en la notación hexadecimal de números.

SISTEMA DE NUMERACIÓN HEXADECIMAL

El sistema de numeración hexadecimal usa la base 16, en contraposición con la base 10 del sistema decimal y la base 2 del binario. Los 16 caracteres diferentes que se necesitan para representar cada dígito hexadecimal pueden elegirse un poco arbitrariamente, pero las convenciones corrientes utilizan los 10 dígitos decimales de 0 a 9 y las letras A, B, C, D, E y F. Los dígitos hexadecimales y su equivalencia decimal se muestran en la Tabla II.

Un dígito hexadecimal aislado puede representar inequívocamente cuatro bits binarios porque hay 16 esquemas diferentes de cuatro dígitos binarios en los números de 0000 a 1111. Éstos también se muestran en la Tabla II. Una impresión de memoria en notación hexadecimal es una representación del contenido exacto del almacenamiento. Sólo se necesita traducir el hexadecimal en binario para saber exactamente los bits que están almacenados. Por ejemplo, supongamos que la computadora está considerando cuatro bytes como una única unidad de datos. El valor interno podría ser.

11000101001000000001011011111101

en números binarios. Este número se representa más fácilmente en hexadecimal, convirtiendo cada grupo de cuatro dígitos binarios en un dígito hexadecimal.

C 5 2 0 1 6 F D

Ahora, un cambio de un bit del número binario no se detecta fácilmente en la forma binaria. ¿Puede encontrar qué bit se cambió abajo?

11000101001000000001111011111101

Una impresión en hexadecimal es

C 5 2 0 1 E F D

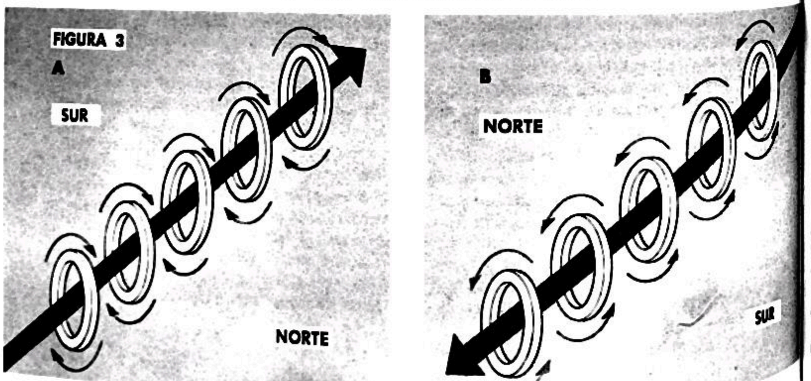
El cambio de 6 (0110) a E (1110) se halla muy fácilmente.

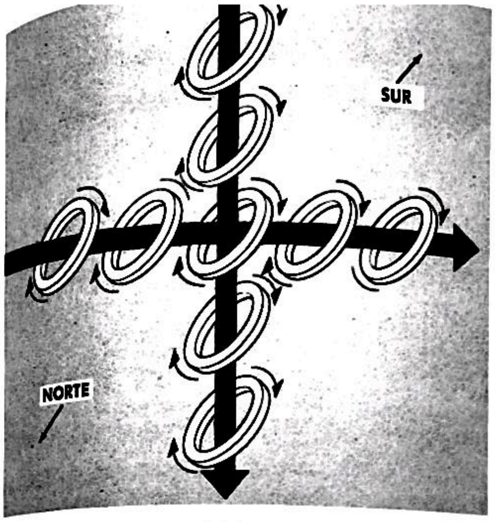
O supongamos que la descarga de un núcleo produce el número hexadecimal siguiente

8 0 1 1

Se puede escribir fácilmente la forma binaria como

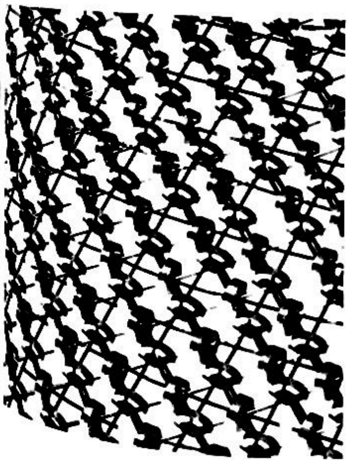
Cuando los núcleos de memoria magnética se enhebran en un cable, el campo magnético generado por una corriente que circula a lo largo de él magnetiza los núcleos en la misma dirección. Las Figuras 3*a* y 3*b* muestran qué sucede cuando la dirección de la corriente es invertida: la dirección del campo se invierte.





La fuerza de la corriente que pasa por un cable se regula de tal modo que sea muy débil para magnetizar un núcleo en forma permanente. Sólo con dos pulsaciones eléctricas que actúen al mismo tiempo a través de dos cables la corriente será lo bastante fuerte para magnetizar un núcleo en forma permanente.

Parte de un plano de núcleos, grandemente ampliados. Cada núcleo puede almacenar un bit de información. Adviértase cómo cada núcleo está ubicado en la intersección de dos cables.



100000000010001

y si este valor se considera como un único entero binario con un bit de signo y 15 bits de datos, el valor del entero decimal es −17.

Las computadoras no operan con números hexadecimales. Ellos son sólo una forma de representar números binarios. Son importantes en las computadoras actuales con memorias organizadas sobre la base de bytes, porque los dígitos hexadecimales representan la cantidad exacta de bits que hay en un byte de la memoria.

Las memorias de núcleo magnético son rápidas. Las computadoras actuales pueden trasferir un byte de la memoria al procesador central en cerca de un microsegundo (una millonésima de segundo). Este "tiempo de trasferencia" se denomina *tiempo de acceso*. Hoy día puede disponerse de computadoras con tiempos de acceso del orden de 100 a 250 nanosegundos (cienmillonésimas de segundo). Las memorias de núcleos son caras; por tanto, muchos sistemas de computadoras tienen una cantidad limitada de aquéllas y las acrecen con otros tipos más baratos, pero también más lentos.

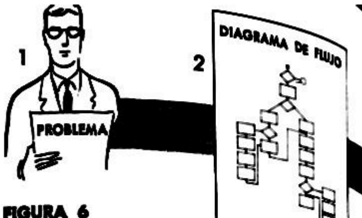
TABLA II

NÚMEROS HEXADECIMALES

DECIMAL	*HEXADECIMAL*	*BINARIO*
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
10	A	1010
11	B	1011
12	C	1100
13	D	1101
14	E	1110
15	F	1111

OPERACIÓN EFICIENTE

La computadora está proyectada para satisfacer las necesidades de computación de una variedad de usuarios. Aun la pequeña computadora descrita en esta parte del artículo podría servir a los estudiantes de un grupo de computación o un curso de matemáticas, o a la oficina administrativa del colegio para ciertas funciones del personal. Aun el bibliotecario puede usarla pasa saber dónde están los libros retirados. Los procedimientos de operación, sin embargo, deben ejecutarse de manera que cada usuario posea un medio sencillo de acceso.



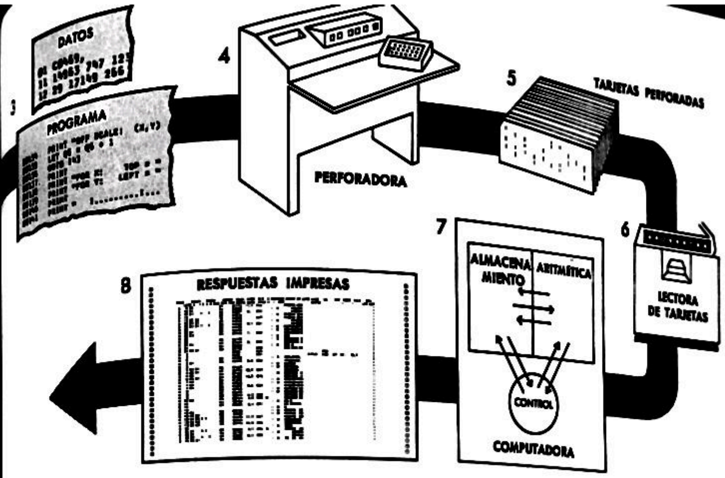


FIGURA 6

Cómo puede operar una pequeña computadora electrónica. El usuario (1) analiza el problema y lo expresa en forma de un diagrama de flujo (2). El usuario traduce entonces los pasos del diagrama de flujo en planteos en lenguaje de programación y prepara los datos para el programa (3). Los datos y los planteos del programa se perforan en tarjetas (4, 5). El usuario entonces se sienta frente a la consola de la computadora, al alcance de los distintos conmutadores y botones. Siguiendo un procedimiento establecido, activa el sistema de computación. La computadora lee las tarjetas en el almacenamiento con el dispositivo lector (6), procesa los datos de acuerdo con el programa (7) e imprime los resultados (8).

Puede considerarse que cada usuario tiene una tarea que realizar, y generalmente quiere completarla en seguida. Un procedimiento es, por tanto, permitir a cada cual hacerse cargo de toda la computadora para procesar el trabajo. El usuario analiza el problema y prepara el programa a fin de efectuar los pasos necesarios para las computaciones. Luego, acciona conmutadores y aprieta los botones apropiados en la consola para que la máquina lea tarjetas, efectúe la computación e imprima el resultado. Sólo una vez terminado el trabajo, trasfiere el control del equipo al próximo usuario (véase la Figura 6).

Éste es el modo más ineficaz de usar la computadora. Puede pasar largo tiempo entre la terminación de una tarea y el comienzo de la próxima, y todo el equipo queda sin funcionar, o bien algunas partes permanecen inactivas por largos lapsos, mientras otras trabajan mucho. Por ejemplo, el dispositivo impresor puede permanecer ocioso durante una larga computación que mantenga la UCP y el almacenamiento muy ocupados.

Si la computadora debe usarse casi enteramente para instrucción, esta experiencia de computación "manual" sirve, por cierto, para eliminar algo del misterio de la computación. La ineficiencia es parte del costo del aprendizaje. En grupos de ingeniería, aun dentro de grandes compañías, una pequeña computadora puede operarse de este modo para la conveniencia individual.

A medida que más usuarios y tareas crecientemente complejas comienzan a presionar sobre el medio de computación, la necesidad de mejores procedimientos de manejo se hace notoria. La magnitud de la tarea confiada a la computadora exige que la máquina esté activa en todo momento, que no permanezca ociosa mientras un usuario está terminando su trabajo y el próximo no tiene aún listo el suyo. El mecanismo impresor no debería quedar inactivo mientras un usuario está haciendo una computación muy larga. Para aumentar la eficiencia así como la flexibilidad, el equipo de computación se pone bajo el control de un procedimiento llamado de varias maneras: programa *monitor*, programa *de control de supervisión*, programa *ejecutivo*, o *sistema de operación*. De un modo u otro, estos programas regulan la operación de la computadora; planean la sucesión de trabajos, asignan a las distintas unidades diferentes tareas y mantienen un registro contable respecto del uso de la máquina. Varían en tamaño, complejidad y funcionamiento, pero básicamente relevan al usuario de operar él mismo la computadora.

El sistema de operación más sencillo se conoce como *procesador por tandas*. Esencialmente, es una versión automatizada del procedimiento de operación manual recién descrito. La obligación de la computadora es aún realizar una sucesión de tareas, cada una de las cuales exige su servicio completo hasta que se concluya. A pesar de que las tareas individuales no se relacionan entre sí, forman una corriente continua de trabajo, interactuando con el sistema de procesador por tandas, donde reside en realidad el control básico.

En el procesamiento por tantas, una serie de tareas, a menudo la tarea de todo un día, se prepara anticipadamente en algún órgano de entrada. Por ejemplo, con la entrada de tarjetas perforadas, los mazos de tarjetas que representan las diferentes tareas por realizar se acumulan en grandes bandejas. Se someten al procesador por tandas, que introduce una tarea por vez, así como el respectivo programa y sus datos. Una vez completado el procesamiento y la salida, el procesador por tandas introduce la próxima tarea. Esto continúa hasta que todas las tareas se han completado.

El usuario ya no ejerce control sobre la computadora. Simplemente, somete una tarea al centro de computación para su procesamiento y luego recoge el producto cuando está completo o cuando le ha sido devuelto. La costosa máquina de computar se mantiene así más activa realizando siempre trabajo útil.

HARDWARE Y SOFTWARE

Los componentes mecánicos o electrónicos de una computadora, y su conjunto, reciben en

inglés el nombre de *hardware*. En cambio, se llama *software* la totalidad de programas escri- tos e instrucciones para computadoras digitales. Tanto los sistemas de *hardware* como los de *software* hoy usados son mucho más complejos que los de la simple computadora recién des- crita, pero los principios son los mismos.

La computadora es una invención reciente en la larga historia de las máquinas. A media- dos del siglo XX salió del laboratorio de inves- tigación de las universidades para convertirse en base de una gran industria y servir en to- das las tareas imaginables y en aplicaciones científicas. Las velocidades internas de compu- tación se redujeron de un segundo por cálculo, a principios de la década de 1940, a menos de un microsegundo en los modelos actuales. Mé- todos de entrada y salida de capacidad limi- tada han sido reemplazados con una variedad de diapositivos casi interminable. Las unidades caras de núcleos de almacenamiento se aumen- tan con discos magnéticos de gran capacidad, pero de tiempo de acceso más lento. La Figura 7 es un diagrama de un conjunto casi com- pleto de unidades que pueden ser parte de un sistema complejo.

Las cintas magnéticas vienen en carretes de 800 metros, capaces de acumular datos en den- sidades de 625 caracteres por centímetro y ser leídos o escritos a velocidades de hasta 180.000 caracteres por segundo. Los cheques de banco se imprimen con caracteres en tinta magné- tica que la computadora puede leer automática- mente. Otros documentos se imprimen con un tipo especial de caracteres que un dispositivo de percepción óptica puede leer. En realidad, el elemento impresor de la computadora puede imprimir con estos caracteres una matriz de salida, por ejemplo una factura, luego enviada al cliente y devuelta por éste con el pago. Una vez agregado el nuevo dato, la factura pasa al dispositivo lector y retorna a la UCP. Estas ma- trices de salida son de uso común en las com- pañías de servicios públicos.

DISPOSITIVOS DE ENTRADA Y SALIDA

Máquinas de escribir, teletipos y otras clases de teclados se usan también como dispositivos de entrada de control remoto si se conectan a la computadora, ya sea por medio de un cable directo, o de un circuito telefónico común. En este caso, gracias a un acoplador telefónico es- pecial, se comunican con un centro de compu- tación desde muchos kilómetros de distancia. Una persona que escriba en Nueva York en una máquina así conectada, puede comunicarse con una computadora de Los Ángeles.

Algunos de los aparatos de teclado están especialmente diseñados para una aplicación particular. Algunos están también combinados con un tubo de rayos catódicos, o TRC, dis- positivo de imagen similar a un aparato de televisión. La salida de la computadora se ex- hibe en el TRC. Los TRC están equipados, asimismo, con "lápices de luz" para convertir- se también en una unidad de entrada. Activando el lápiz de luz y apuntándolo a lugares espe- ciales en la superficie del TRC, se inicia el proceso de computación. Los dispositivos de TRC más completos exhibirán gráficas en tres dimensiones (o en dos) y en color. Para la entrada, se dibuja en la superficie del tubo.

El teléfono mismo, especialmente el modelo con teclas en vez de disco, se está usando co- mo dispositivo de entrada y salida. Sus teclas trasmiten a la computadora tanto códigos nu- méricos como alfabéticos. El procesador central está equipado para contestar el teléfono, tal como un conmutador. La salida de la compu- dora se traduce en señales que luego se tras- miten nuevamente al aparato en forma de voz. Así, uno puede llamar a la computadora, enviar mensajes desde el panel del teléfono y recibir una contestación en voz audible.

Otro dispositivo de salida es el graficador, que puede realizar gráficas de datos generados por computadora. Las gráficas pueden ser de puntos o lineales de gran complejidad y exac- titud. Algunos diseños de gráficas tienen valor artístico.

ALMACENAMIENTO DE DISCO

El caro, pero rápido, almacenamiento de núcleo está siendo acrecido por unidades de almacenamiento de discos magnéticos que tie- nen mayor capacidad, pero tiempo de acceso más lento a los datos. Las unidades de almace- namiento de disco permiten que grandes canti- dades de datos, y aun de programas, se almace- nen en el sistema central de la computadora para el acceso de la UCP. Algunos sistemas llegan a tener unos 1.000.000.000 de bytes o caracteres de dicho almacenamiento.

La unidad de almacenamiento de disco es

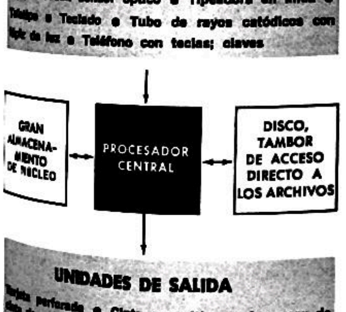
un plato, o disco giratorio de metal que se re- viste con óxido de hierro para que pueda ser magnetizado con puntos que representen el 0 ó el 1 en los bits de un byte. Estos discos se montan sobre un eje que gira a 2.500 revolu- ciones por minuto. Se apilan con varios otros sobre un eje único, pero dejando espacio entre ellos. Un brazo con dos cabezas, una de lectura y otra de escritura, se sitúa en tal forma que se mueva entre las superficies de dos discos. Una cabeza se encarga de la cara inferior del disco superior, y la otra de la cara superior del disco inferior. Un conjunto completo de dichos bra- zos se mueve al unísono para llegar a todas las caras al mismo tiempo, ubicándose en la misma pista concéntrica de cada cara y movién- dose de una pista a otra. Estos brazos múltiples permiten trasferir datos en milisegundos a la memoria de núcleos para el procesamiento.

En el centro de computación se usan gran- des archivos de discos como almacenamiento permanente, para datos de inmediato acceso, en contraposición con los archivos de cintas, que son accesibles sólo después que las cintas se han montado. Los datos de los archivos de discos son también accesibles sin necesidad de buscar sucesivamente en el archivo co- mo con la cinta. Los brazos pueden colo- carse adecuadamente en la pista donde están grabados los datos requeridos. Por esta razón, dichas unidades de almacenamiento de disco se denominan unidades de *acceso directo*.

El diagrama muestra la gran variedad de unidades de en- trada y salida accesibles a los usuarios de computadoras.

FIGURA 7

UN SISTEMA DE COMPUTACIÓN COMPLEJO



SISTEMAS DE OPERACIÓN

El desarrollo de dichos adelantos en los dis- positivos mecánicos y electrónicos de las com- putadoras (hardware), ha sido paralelo al desa- rrollo de sistemas de operación adecuados para manejarlos (software). Veamos, como ejemplo, un sistema de reserva de pasajes aéreos.

Una línea aérea debe tener cientos, y quizá miles, de dispositivos de teclados de control re- moto en sus oficinas de pasajes. Estos teclados enlazan con un sistema central de computación a través de un extenso conjunto de líneas de comunicación. El pedido de asiento para deter- minado vuelo se envía directamente a la com- putadora central, donde hay datos sobre todos los vuelos preparados con varios meses de an- ticipación, en un gran dispositivo de almace- namiento de acceso directo. Cuando el pedido es aceptado por el agente vendedor de pasajes, la UCP del sistema de computación averigua si el asiento está disponible gracias a la infor- mación de vuelos almacenada en el disco. La salida, que será la confirmación o rechazo de la reserva, se trasmite al agente. El cliente puede entonces ir a una oficina de pasajes, o tratar el asunto por teléfono con el agente. Sea como fuere, este último está en comunicación directa con la computadora central para cum- plir el trámite de venta de pasajes.

Una exigencia para el desarrollo de dicho sistema de reservas es una red de comunicación por medio de la cual los miles de terminales puedan trasmitir y recibir datos. Otra es la disponibilidad de grandes unidades de alma- cenamiento de disco donde se guarden los datos de todos los vuelos. Esto incluye no sólo in- formación sobre éstos, sino también sobre pa- sajeros, como nombres, clase de servicio, nú- mero telefónico local, cómo pagarán el pasaje, etc. La tercera exigencia es un buen sistema de operación, o *software*, para poner en marcha el sistema de computación.

El sistema de operación debe responder al pedido de servicio de un agente. Puede nece- sitar la atención de varios terminales simultá- neamente. Debe decodificar el mensaje recibido

inglés el nombre de *hardware*. En cambio, se llama *software* la totalidad de programas escritos e instrucciones para computadoras digitales. Tanto los sistemas de *hardware* como los de *software* hoy usados son mucho más complejos que los de la simple computadora recién descrita, pero los principios son los mismos.

La computadora es una invención reciente en la larga historia de las máquinas. A mediados del siglo XX salió del laboratorio de investigación de las universidades para convertirse en base de una gran industria y servir en todas las tareas imaginables y en aplicaciones científicas. Las velocidades internas de computación se redujeron de un segundo por cálculo, a principios de la década de 1940, a menos de un microsegundo en los modelos actuales. Métodos de entrada y salida de capacidad limitada han sido reemplazados con una variedad de dispositivos casi interminable. Las unidades caras de núcleos de almacenamiento se aumentan con discos magnéticos de gran capacidad, pero de tiempo de acceso más lento. La Figura 7 es un diagrama de un conjunto casi completo de unidades que pueden ser parte de un sistema complejo.

Las cintas magnéticas vienen en carretes de 800 metros, capaces de acumular datos en densidades de 625 caracteres por centímetro y ser leídos o escritos a velocidades de hasta 180.000 caracteres por segundo. Los cheques de banco se imprimen con caracteres en tinta magnética que la computadora puede leer automáticamente. Otros documentos se imprimen con un tipo especial de caracteres que un dispositivo de percepción óptica puede leer. En realidad, el elemento impresor de la computadora puede imprimir con estos caracteres una matriz de salida, por ejemplo una factura, luego enviada al cliente y devuelta por éste con el pago. Una vez agregado el nuevo dato, la factura pasa al dispositivo lector y retorna a la UCP. Estas matrices de salida son de uso común en las compañías de servicios públicos.

DISPOSITIVOS DE ENTRADA Y SALIDA

Máquinas de escribir, teletipos y otras clases de teclados se usan también como dispositivos de entrada de control remoto si se conectan a la computadora, ya sea por medio de un cable directo, o de un circuito telefónico común. En este caso, gracias a un acoplador telefónico especial, se comunican con un centro de computación desde muchos kilómetros de distancia. Una persona que escriba en Nueva York en una máquina así conectada, puede comunicarse con una computadora de Los Ángeles.

Algunos de los aparatos de teclado están especialmente diseñados para una aplicación particular. Algunos están también combinados con un tubo de rayos catódicos, o TRC, dispositivo de imagen similar a un aparato de televisión. La salida de la computadora se exhibe en el TRC. Los TRC están equipados, asimismo, con "lápices de luz" para convertirse también en una unidad de entrada. Activando el lápiz de luz y apuntándolo a lugares especiales en la superficie del TRC, se inicia el proceso de computación. Los dispositivos de TRC más completos exhibirán gráficas en tres dimensiones (o en dos) y en color. Para la entrada, se dibuja en la superficie del tubo.

El teléfono mismo, especialmente el modelo con teclas en vez de disco, se está usando como dispositivo de entrada y salida. Sus teclas trasmiten a la computadora tanto códigos numéricos como alfabéticos. El procesador central está equipado para contestar el teléfono, tal como un conmutador. La salida de la computadora se traduce en señales que luego se trasmiten nuevamente al aparato en forma de voz. Así, uno puede llamar a la computadora, enviar mensajes desde el panel del teléfono y recibir una contestación en voz audible.

Otro dispositivo de salida es el graficador, que puede realizar gráficas de datos generados por computadora. Las gráficas pueden ser de puntos o lineales de gran complejidad y exactitud. Algunos diseños de gráficas tienen valor artístico.

ALMACENAMIENTO DE DISCO

El caro, pero rápido, almacenamiento de núcleo está siendo acrecido por unidades de almacenamiento de discos magnéticos que tienen mayor capacidad, pero tiempo de acceso más lento a los datos. Las unidades de almacenamiento de disco permiten que grandes cantidades de datos, y aun de programas, se almacenen en el sistema central de la computadora para el acceso de la UCP. Algunos sistemas llegan a tener unos 1.000.000.000 de bytes o caracteres de dicho almacenamiento.

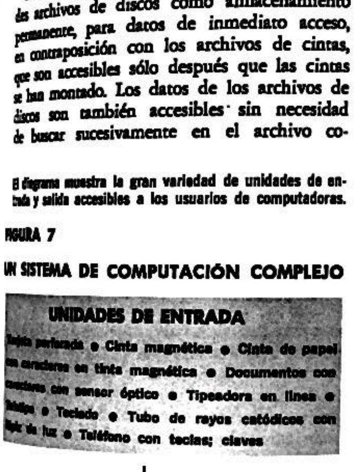
La unidad de almacenamiento de disco es un plato, o disco giratorio de metal que se reviste con óxido de hierro para que pueda ser magnetizado con puntos que representen el 0 ó el 1 en los bits de un byte. Estos discos se montan sobre un eje que gira a 2.500 revoluciones por minuto. Se apilan con varios otros sobre un eje único, pero dejando espacio entre ellos. Un brazo con dos cabezas, una de lectura y otra de escritura, se sitúa en tal forma que se mueva entre las superficies de dos discos. Una cabeza se encarga de la cara inferior del disco superior, y la otra de la cara superior del disco inferior. Un conjunto completo de dichos brazos se mueve al unísono para llegar a todas las caras al mismo tiempo, ubicándose en la misma pista concéntrica de cada cara y moviéndose de una pista a otra. Estos brazos múltiples permiten trasferir datos en milisegundos a la memoria de núcleos para el procesamiento.

En el centro de computación se usan grandes archivos de discos como almacenamiento permanente, para datos de inmediato acceso, en contraposición con los archivos de cintas, que son accesibles sólo después que las cintas se han montado. Los datos de los archivos de discos son también accesibles sin necesidad de buscar sucesivamente en el archivo como con la cinta. Los brazos pueden colocarse adecuadamente en la pista donde están grabados los datos requeridos. Por esta razón, dichas unidades de almacenamiento de disco se denominan unidades de *acceso directo*.

El diagrama muestra la gran variedad de unidades de entrada y salida accesibles a los usuarios de computadoras.

FIGURA 7

UN SISTEMA DE COMPUTACIÓN COMPLEJO



SISTEMAS DE OPERACIÓN

El desarrollo de dichos adelantos en los dispositivos mecánicos y electrónicos de las computadoras (hardware), ha sido paralelo al desarrollo de sistemas de operación adecuados para manejarlos (software). Veamos, como ejemplo, un sistema de reserva de pasajes aéreos.

Una línea aérea debe tener cientos, y quizá miles, de dispositivos de teclados de control remoto en sus oficinas de pasajes. Estos teclados enlazan con un sistema central de computación a través de un extenso conjunto de líneas de comunicación. El pedido de asiento para determinado vuelo se envía directamente a la computadora central, donde hay datos sobre todos los vuelos preparados con varios meses de anticipación, en un gran dispositivo de almacenamiento de acceso directo. Cuando el pedido es aceptado por el agente vendedor de pasajes, la UCP del sistema de computación averigua si el asiento está disponible gracias a la información de vuelos almacenada en el disco. La salida, que será la confirmación o rechazo de la reserva, se trasmite al agente. El cliente puede entonces ir a una oficina de pasajes, o tratar el asunto por teléfono con el agente. Sea como fuere, este último está en comunicación directa con la computadora central para cumplir el trámite de venta de pasajes.

Una exigencia para el desarrollo de dicho sistema de reservas es una red de comunicación por medio de la cual los miles de terminales puedan trasmitir y recibir datos. Otra es la disponibilidad de grandes unidades de almacenamiento de disco donde se guarden los datos de todos los vuelos. Esto incluye no sólo información sobre éstos, sino también sobre pasajeros, como nombres, clase de servicio, número telefónico local, cómo pagarán el pasaje, etc. La tercera exigencia es un buen sistema de operación, o *software*, para poner en marcha el sistema de computación.

El sistema de operación debe responder al pedido de servicio de un agente. Puede necesitar la atención de varios terminales simultáneamente. Debe decodificar el mensaje recibido

y ordenar a la computadora el comienzo del procesamiento de programas que leerán los datos correctos del vuelo en el disco, ingresarán la información sobre el pasajero y almacenarán nuevamente en aquél la información de vuelo actualizada, y luego controlarán la trasmisión de la confirmación del pasajero al agente. El sistema de operación es un elaborado mecanismo de control en esta reserva por líneas. Sus tareas ya no son tan sencillas como controlar una tanda de trabajos independientes.

En un contexto diferente, podemos imaginar una fábrica que tiene un sistema de control de inventario similar al sistema de reservas de la línea aérea, excepto que se piden datos sobre piezas y productos, no sobre asientos y pasajeros. Además de controlar el inventario, la computadora central también atiende a un grupo de ingenieros que tiene que hacer pequeñas tareas de computación. Éstos pueden usar un terminal remoto para someter a procesamiento sus programas y datos. La oficina administrativa de la planta usa la computadora para solicitudes de personal, procesamiento de la lista de pagos y proyectos financieros.

Estas aplicaciones representan una variedad de tipos de servicio. El control del inventario representa una actividad que requiere servicio inmediato para actualizar los archivos del mismo. Algunos ingenieros necesitan estar en contacto directo con la computadora central, mientras otros someterán tareas que pueden realizarse en la noche para preparar el trabajo del próximo día. Ésa es una forma de procesamiento en tanda. La oficina de pagos libra cheques por semana o por mes. No necesita comunicación continua con la computadora; le basta un tipo corriente de procesamiento en tanda.

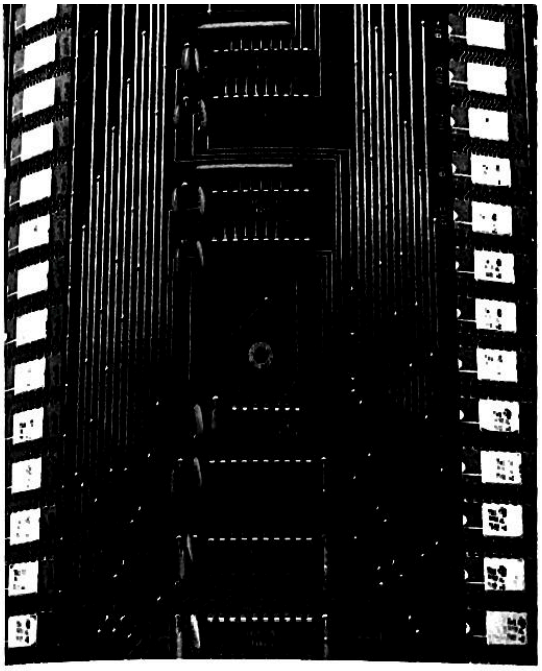
El sistema de operación debe controlar ahora estos distintos niveles de actividad. Multiprogramación es el nombre dado a un tipo de sistema. La clave de la operación es la partición del almacenamiento de núcleo a fin de que cada usuario tenga una parte del mismo para su propio uso. El sistema de operación mantiene el programa de cada usuario operando simultáneamente; la computadora está trabajando en diferentes tareas al mismo tiempo. Una impresión de cheques de pago de baja prioridad puede continuar sin interrupción durante todo el día, mientras que los trabajos de ingeniería se trasfieren dentro del almacenamiento y fuera de él según la demanda. Un pedido de servicio del sistema de control de inventario tiene prioridad sobre cualquier otro, por eso otro programa tendrá que esperar mientras el pedido se satisface. Por ejemplo, la imprenta ssupenderá la impresión de cheques por unos pocos milisegundos mientras se examina el archivo del inventario de repuestos.

TIEMPO COMPARTIDO

La mayor mejora lograda hace poco en los sistemas que rigen las operaciones de computación es el *tiempo compartido*. Un gran número de usuarios se ubica ante sus propios terminales (máquina de escribir, teletipo, teclado o dispositivo de proyección) conectados con la computadora central. Cada uno le somete un problema. Uno puede estar escribiendo un programa; un segundo, corrigiendo el suyo; un tercero, enviando datos a otro; un cuarto, leyendo su salida, etc. Cada uno parece controlar la computadora como si fuera el único usuario, como en la operación de calculadoras manuales; pero, en realidad, la está compartiendo con doscientas o trescientas personas más.

La posibilidad de compartir la computadora se debe a la rapidez con que ésta procesa los datos en relación con la velocidad con que la gente puede usar un dispositivo terminal. Por ejemplo, mientras un usuario pierde cinco segundos en ingresar algunos datos (pruebe y vea cuántos logra escribir a máquina en cinco segundos), la UCP de una computadora moderna puede ejecutar hasta un millón de computaciones, o sea, 10.000 para cada uno de diez usuarios, y sin embargo no retarda a la persona que está ingresando otros datos. Cuando un usuario termina, la UCP procesa sus datos mientras alguien está examinando algún resultado o pensando qué hará luego.

En un sistema de tiempo compartido, el usuario y la computadora están interactuando entre sí en lo que algunos llaman un coloquio. El usuario hace algo, la computadora responde. Aquél hace algo más, y la computadora responde nuevamente. Esto continúa en una suerte de diálogo entre el hombre y la máquina. Los terminales remotos en un sistema de tiempo compartido pueden estar a muchos kilómetros de la computadora, en cualquier parte donde pueda disponerse de un teléfono común. Con un acoplador acústico (un dispositivo



Hewlett Packard

La unidad primaria de almacenamiento de una computadora se denomina memoria: recibe los datos, los almacena y los suministra para su procesamiento. Arriba esta el tablero de la compleja memoria de una computadora Hewlett Packard 3000 Serie 2.

que permite a una máquina de escribir o un sistema de teclado intercomunicarse con un receptor telefónico común) uno puede discar y comunicarse con el sistema. En realidad, es factible comprar dicho servicio de tiempo compartido a cualquiera de las varias firmas de computadoras, pagando por el uso de la máquina según las necesidades y los costos del teléfono. Por ejemplo, sin poseer siquiera una máquina, los alumnos de un colegio secundario local podrían usar una computadora si contratan los servicios de tiempo compartido. En realidad, varios colegios podrían usar una computadora central para sus necesidades locales en dichos términos. Un alumno puede estar resolviendo una ecuación desde el terminal de su colegio mientras otro, de un colegio distinto, está aún escribiendo su programa.

INFINITO NÚMERO DE USOS

Con el desarrollo de los más avanzados sistemas de computación y operación, la lista de las aplicaciones comunes de las computadoras ha crecido casi sin límite. Procesan ahora boletas de votos que se "marcan" en alguna forma legible por la máquina, tal como perforar una tarjeta. También procesan cheques bancarios cuyos números de cuenta están preimpresos en caracteres magnéticos para entrada automática. Muchos en la comunidad bancaria están también preparando una sociedad "sin cheques". Cuando un cliente compre algo en una tienda, la compra se registraría y los datos irían directamente al centro de computación del banco que sirve al establecimiento. Se haría un depósito en la cuenta de la tienda por el importe, porque los fondos se retirarían de la cuenta del cliente en otro banco mediante sus sistemas de computadoras. No sólo no habría dinero implicado; ni siquiera habría cheques.

La política está influida por grandes archivos de datos sobre las características de los votantes y la composición de los distritos electorales. Los datos se actualizan basándose en encuestas

Sperry Rand Corporation

Intel Corp.

de opinión pública para mantener a los líde-res políticos informados sobre las reacciones del pueblo ante los eventos y estrategias de las campañas.

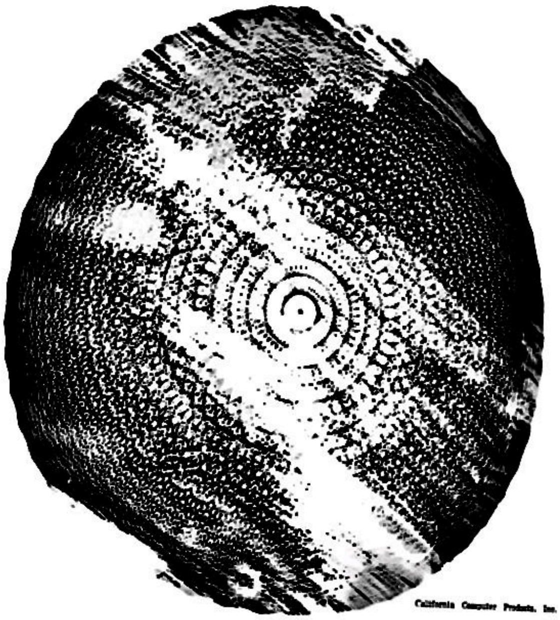
Pueden almacenarse datos de los pacientes en un hospital, para que los médicos reciban más rápidamente las historias clínicas. También es posible analizar a los pacientes conectando a la computadora dispositivos que miden sus reacciones, a fin de que los datos sean inmediatamente asequibles al médico para mejorar su servicio.

La lista es indefinida y, en cierto modo, estamos nuevamente al principio. El motivo del desarrollo de la computadora electrónica a comienzos de la década del cuarenta fue la necesidad de computar las trayectorias de proyectiles para nuevas armas. Actualmente, las computadoras que surgieron del laboratorio computan las órbitas de los satélites de la Tierra y la trayectoria de vuelo de las naves espaciales que alunizan. Puede muy bien decirse que dichos adelantos científicos no hubieran sido posibles sin la moderna computadora electrónica.

Tanto lo bueno como lo malo pertenecen al vasto potencial de la computadora. Las grandes unidades de almacenamiento que sirven a las líneas aéreas y a los bancos y fábricas también permiten la recolección y almacenamiento centralizado de datos personales. En los Estados Unidos, el número de seguridad social está volviéndose la clave para identificar a sus ciudadanos. Existen ya bancos de datos de crédito, y de personal, de la salud y de licencias de conducir. Lo privado se está volviendo un tema de discusión, y los debates sobre el buen y mal uso de los datos continuará.

Esta sola razón debiera bastar para que una persona informada adquiriera una comprensión básica de las computadoras, pues sólo a partir de dicha comprensión será capaz de contribuir a maximizar los beneficios y minimizar el perjuicio de la más fascinante de todas las máquinas: la computadora electrónica.

Izquierda, arriba: varias personas utilizan una computadora UNIVAC 9070. Izquierda, al medio: técnicos que ejecutan sobre un plano el intrincado diseño de un chip para microprocesador. Los microprocesadores son minúsculos dispositivos de silicio que contienen los complejos circuitos de una unidad central de procesamiento de una microcomputadora. Izquierda, abajo: trabajadores con microscopios arman los chips en paquetes.



California Computer Products, Inc.

El arte pictórico matemáticamente programado es una emocionante creación de la época de la computación. El artista introduce ecuaciones y demás información en una computadora para lograr un dibujo. Este, de un erizo de mar, recuerda una vidriera de colores

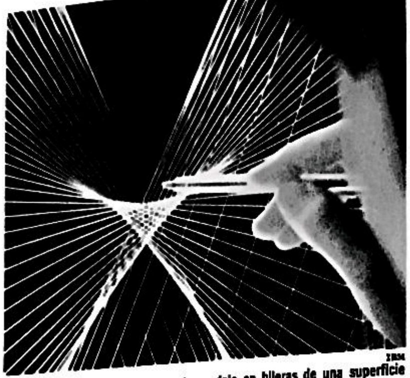
PROGRAMACIÓN DE COMPUTADORAS

Para muchos, la computadora electrónica es un tipo de robot sobrehumano que puede resolver cálculos a velocidad fulmínea, tomar decisiones en fracciones de segundo y descender hombres en la Luna. Es cierto: hay computadoras capaces de hacer varios millones de sumas en un segundo, tomar decisiones y guiar un módulo lunar a un alunizaje suave. Pero la computadora no es sobrehumana; tampoco es un robot, porque no puede realizar ninguna de estas cosas por sí sola.

A cualquiera de las computadoras que existen actualmente debe decírsele qué hacer; debe dársele un conjunto de instrucciones. Estas instrucciones se denominan *programa.* La escritura de estos instrumentos se denomina *programación de computadoras.*

La computadora no puede pensar, pero sí realizar procesos que un ser humano piensa y programa para ella.

La computadora sólo es capaz de cumplir un número limitado de instrucciones. Con este limitado conjunto, el hombre debe describir procesos a veces muy complejos. Tales descrip-ciones a menudo someten sus facultades men-tales a un esfuerzo máximo. Esto contribuye a que sea divertido aprender programación de computadoras. El entretenimiento realmente proviene de dos fuentes. Una es idear un pro-ceso que creemos correcto y que la computa-



Los planos para este modelo en hileras de una superficie generada matemáticamente los dibujó una computadora.

dora debería ser capaz de seguir. La otra es comprobar si el procedimiento es de verdad correcto y si la computadora puede llevarlo a cabo.

El proceso que presuntamente debe efectuar la computadora se describe en lenguaje de programación. En la Tabla 1 se muestran cuatro ejemplos de programación. Cada programa indicado cumple una tarea muy sencilla: lee dos números, los suma e imprime el total. El programa final, en lenguaje de máquina, consta de un conjunto de números binarios, 0 y 1, que la unidad central de procesamiento de la computadora puede interpretar directamente y ejecutar. El programa final es ininteligible para cualquiera que no conozca bastante acerca de cómo está constituida la computadora y cómo representa los datos e instrucciones en el sistema de numeración binaria. En contraposición con el lenguaje de máquina, el programa LP/1 es casi autoexplicativo. LP/1 es una abreviatura de Lenguaje de Programación Uno.

En la Tabla 1 se puede elegir uno de los dos lenguajes –LP/1 o FORTRAN– que utilizará el programador. En esta tabla, estamos sumando dos números e imprimiendo el resultado. LP/1 y FORTRAN son ejemplos de lenguajes a nivel del usuario.

En algunas computadoras, las instrucciones contenidas en FORTRAN o LP/1 deben llevarse a un nivel más adecuado para ellas. Un lenguaje en este nivel intermedio se llama *lenguaje de agrupamiento.*

Finalmente, las instrucciones deben codificarse en números binarios: el lenguaje que la computadora entiende. Esta sucesión de números binarios se denomina *lenguaje de máquina.*

Cada programa de muestra hace a su manera que la computadora siga una serie de pasos para completar la tarea dada. El lenguaje de máquina está al nivel de la computadora misma. Los otros se trasladan luego de este nivel a un lenguaje que pueda leerse y comprenderse. En cualquiera de estos y otros lenguajes de programación puede establecerse con precisión el proceso que usará la computadora.

LENGUAJE DE MÁQUINA

Dos dispositivos importantes para un conocimiento de la programación son la unidad central de procesamiento, o UCP, y la unidad de almacenamiento, o memoria.

La UCP es el dispositivo que ejecuta las instrucciones del calculista y controla toda la operación de la computadora. El dispositivo de almacenamiento es el lugar donde se guardan los datos e instrucciones para el uso de la UCP.

En su forma más sencilla, una instrucción del calculista que puede realizarse con la UCP consta de dos partes. Una es un código de operación que da órdenes. El código de operación hace que la UCP establezca las conexiones electrónicas necesarias para ejecutar la orden.

TABLA I

Esto es LP/1	TOME LOS DATOS (N, M); K = M + N; COLOQUE DATOS (K);	
Esto es FORTRAN	LEA (2,10) N,M	
	10 FORMAT (14, 14)	
	K = M + N	
	IMPRIMA (3, 11) K	
	11 FORMAT (15)	
Esto es lenguaje simbólico	R	N
	R	M
	CLA	N
	SUME	M
	ALMACENE	K
	IMPRIMA	K
Esto es lenguaje de máquina	1100000000000001	
	1100000000000010	
	1000000000000001	
	1001000000000010	
	1110000000000011	
	1101000000000011	

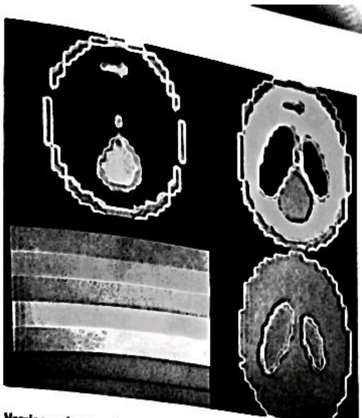
La otra parte es un sector de la memoria donde se almacenan los datos sobre los que se debe actuar. La parte de operación responde a la pregunta: "¿Qué haré?" La otra, a la pregunta: "¿Dónde están los datos?" Las instrucciones en lenguaje de máquina se codifican en una forma que se ha proyectado para una computadora en particular.

Generalmente, está en forma binaria.

Por ejemplo, el programa en lenguaje de máquina mostrado tiene seis instrucciones. Cada instrucción es un número binario de 16 dígitos. La UCP de la computadora toma este número y lo decodifica en dos partes. Los primeros cuatro dígitos binarios son el código de operación. Los 12 restantes son la ubicación de los datos en el almacenamiento. Programar la computadora en lenguaje de máquina significa que debe conocerse la notación binaria para cada código de operación y también cómo expresar las ubicaciones de almacenamiento por medio de números binarios. Puede haber más de 200 códigos de operación en una computadora grande.

El ejemplo de programa puede decodificarse como en la Tabla II. El acumulador que se menciona es simplemente un almacenamiento temporario. Los primeros cuatro dígitos binarios de la instrucción identifican inequívocamente una operación muy sencilla que la UCP puede llevar a cabo o ejecutar. Los 12 dígitos restantes son la ubicación binaria del lugar en que está la memoria donde los datos deben almacenarse o de donde deben tomarse.

Cada instrucción en lenguaje de máquina indica a la computadora que realice sólo una acción muy pequeña. Ésta puede ser extraer un número de la unidad de almacenamiento y ubicado en el acumulador de la UCP, donde puede usarse en una operación aritmética. Puede ser simplemente almacenar un número en la me-



Versiones de computadora de los rayos X de un cerebro. Las bandas continuas son los colores elegidos por un operador de computadora.

moria, o comparar dos números para determinar si son iguales. La computadora electrónica obtiene su poder y capacidad para resolver problemas complejos de la formidable velocidad con que ejecuta estas operaciones elementales.

Un proceso complejo puede necesitar un millón de tales pequeños pasos de instrucciones en un programa a seguir por la computadora. Si cada instrucción puede ejecutarse en una millonésima de segundo (un microsegundo), llevará sólo un segundo completar todo el programa. Las computadoras son capaces de ejecutar instrucciones en mucho menos de un microsegundo; por tanto, procesos muy complejos representados por muchos millones de instrucciones se ejecutan en minutos. Los pequeños procesos que necesitan sólo unos pocos cientos o miles de instrucciones se ejecutan en un milisegundo (milésima de segundo) o menos.

La programación del lenguaje de máquina es un arte en extinción porque está muy ligado al proyecto de la máquina y a sistemas de numeración que no son fácilmente manejables.

TABLA II

Código de operación		Ubicación en el almacenamiento	
1100	lea un valor	000000000001	en ubicación 1
1100	lea un valor	000000000010	en ubicación 2
1000	despeje y sume un valor	000000000001	de la ubicación 1 al acumulador de la UCP
1001	sume un valor	000000000010	de la ubicación 2 al acumulador
1110	almacene el valor del acumulador	000000000011	en la ubicación 3
1101	imprima el valor	000000000011	almacenado en la ubicación 3

Es, sin embargo, la única forma de programa que la UCP puede ejecutar. Un programa escrito en cualquier otra forma debe convertirse a lenguaje de máquina para que una computadora pueda cumplirlo.

LENGUAJE UNIFICADO O DE AGRUPAMIENTO

Un paso aparte de la UCP es la programación en forma simbólica, o lenguaje unificado. En lenguaje unificado, un código de operación se representa por una letra o serie de letras que se refieren a la operación que debe llevarse a cabo. Así, L puede representar LEA; DES, DESPEJE Y SUME EN EL ACUMULADOR; ALM, ALMACENE EL CONTENIDO DEL ACUMULADOR EN LA MEMORIA, e IM, IMPRIMA EL CONTENIDO DE UNA UBICACION DE MEMORIA. Las ubicaciones de datos en el almacenamiento se dan también en letras simbólicas, tales como N, M y K.

El programa de muestra en lenguaje unificado en la Tabla I sigue al programa en lenguaje de máquina orden por orden. Por cada instrucción en lenguaje de máquina aparece una en lenguaje unificado. Como programadores, estamos ahora familiarizados con los códigos simbólicos de operación y los nombres simbólicos que podemos idear para las ubicaciones de almacenamiento, con referencia al problema.

La UCP no puede ejecutar instrucciones de programación simbólicas. Éstas deben traducirse a lenguaje de máquina. Esto lo hace automáticamente la computadora, usando un programa denominado unificador. El programa unificador está en lenguaje de máquina y se suministra con cada computadora cuando se la compra o alquila.

Para muchos que desean resolver problemas en una computadora electrónica, el lenguaje unificado es también demasiado tedioso y detallado. Mucho esfuerzo, por tanto, se ha empleado en idear lenguajes de alto nivel que son luego eliminados de la máquina.

LENGUAJES DE ALTO NIVEL

Los lenguajes de programación de alto nivel permiten escribir las instrucciones de manera similar a la que emplearíamos en el contexto común del problema. Los lenguajes para matemáticos permiten la escritura de fórmulas. Aquellos para asuntos comerciales subrayan las descripciones de archivos de datos comerciales y tipos de operaciones para control de inventarios y liquidaciones de pagos. Se han ideado muchos lenguajes para programar una computadora.

El "paseo" de una computadora alrededor de la Universidad de Cornell, en Nueva York. Arriba: Museo de Arte Johnson, en blanco, "visto" antes de construirse. A la derecha: a medida que la Biblioteca Uris se acerca, la perspectiva cambia. En realidad, la torre fue reemplazada hace años.

Algunos de estos lenguajes son ALGOL, BASIC, COBOL, FORTRAN, GPSS, JOSS, LISP, LP/1, SIMSCRIPT y SNOBOL. Cada uno está destinado a facilitar el trabajo de programación.

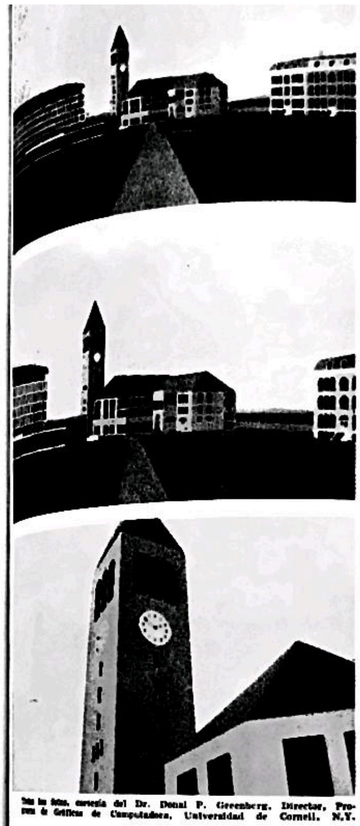
FORTRAN (de FORmula TRANslation = traducción de fórmulas) es un lenguaje muy popular y ampliamente usado. En la Tabla I se exhibe un programa de muestra para una versión de FORTRAN. También se muestra un ejemplo del uso de LP/1. Cada lenguaje tiene sus propias reglas para nombrar variables y escribir planteos para la lectura de datos, realizar operaciones aritméticas e imprimir respuestas. Aprender a programar en un lenguaje es como aprender cualquier lenguaje, excepto que las reglas a veces son más estrictas que las de un lenguaje normal como el castellano.

Los distintos lenguajes de programación, sistemas de programación y programas que se desarrollan en torno de las computadoras se denominan *software* para distinguirlos de las unidades físicas de la computadora, o el *hardware*.

¿QUÉ ES UN ALGORITMO?

Programar en un lenguaje de computadora no es la parte más interesante ni la más difícil en el manejo de una computadora electrónica. La tarea más difícil y más interesante es definir el problema, o sea, idear el proceso que la computadora debe seguir.

Los pasos detallados que deben seguirse se denominan algoritmo. Es el plan que se propone como correcto para el problema y que la computadora debe cumplir. Una vez escrito, el algoritmo puede ser traducido por alguien que conozca el lenguaje de programación, aunque no tenga conocimiento del problema. No obstante, programar su propio problema puede ser parte del entretenimiento.



Todas las fotos, cortesía del Dr. Donal P. Greenberg, Director, Programa de Gráficas de Computadora, Universidad de Cornell, N.Y.

Supongamos, por ejemplo, que, como tarea para realizar en casa, tenemos un problema algebraico que requiere calcular una raíz de una ecuación cuadrática. Dichas ecuaciones son de uso común en los cursos de álgebra. Para resolverlas podríamos querer usar la computadora, una vez que hemos aprendido cómo programar. No nos preocuparemos por un lenguaje de programación más que para decir que debemos describir en detalle los pasos individuales que deben ejecutarse y el orden de ejecución de esos pasos.

La raíz de una ecuación cuadrática se expresa por:

$$\text{raíz} = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

Cuando programamos la solución para esta computación, estamos ideando simplemente un procedimiento computacional, o algoritmo, que dirigirá la computadora. Pensemos cómo explicaríamos a alguien no familiarizado en absoluto con el álgebra, pero con conocimientos de aritmética, cómo debería proceder para resolver esta ecuación. Le daríamos instrucciones paso por paso. Tales instrucciones podrían ser:

1. Lea los valores para A, B y C.
2. Eleve B al cuadrado y escríbalo.
3. Multiplique A por 4 y escríbalo.
4. Multiplique 4A (paso 3) por C y regístrelo.
5. Reste 4AC (paso 4) de B al cuadrado (paso 2).
6. Saque la raíz cuadrada de la diferencia $B^2 - 4AC$.
7. Sume —B a la raíz cuadrada, y consérvelo.
8. Multiplique A por 2 y consérvelo.
9. Divida el resultado del paso 7 por 2A (paso 8).
10. Escriba el resultado del paso 9 como la raíz.

Este procedimiento por pasos conducirá a la computación del valor de la raíz. Se subdivide en pasos pequeños y elementales del tipo que la computadora puede seguir.

Puede haber un punto débil en este procedimiento. ¿Cómo obtenemos la raíz cuadrada en el paso 6? ¿La buscamos en una tabla? ¿Tomamos una regla de cálculo para computarla? ¿Nos dirigimos a otro libro y aprendemos cómo computar una raíz cuadrada a mano? Puede ser necesario subdividir más este paso.

En lenguaje de máquina o unificado se escribirían instrucciones individuales para cada uno de los pequeños pasos mencionados, más otros adicionales para computar la raíz cuadrada. El programa completo a este nivel podría necesitar de 50 a 75 instrucciones. Empero, en lenguaje de nivel "más alto" tal como LP/1, se necesitan sólo 3 planteos.

UN PROGRAMA LP/1 PARA NUESTRO PROBLEMA

Tres planteos en LP/1 efectúan la parte de la raíz cuadrada de la computación: uno para

leer los valores de los datos, uno para computar la raíz y uno para imprimir los resultados. No se requiere una explicación adicional de cómo hallar una raíz cuadrada porque todos los lenguajes de programación definidos para resolver problemas matemáticos incluyen las funciones básicas implícitas. Para computar una raíz cuadrada escribimos un nombre apropiado, como RACU, y la computadora automáticamente realizará los pasos necesarios para computar la raíz cuadrada. Por eso, RACU es el nombre de un conjunto preparado de instrucciones en lenguaje de máquina.

Los símbolos de programación corrientes para aritmética son:

+ para la suma
— para la resta
• para la multiplicación
/ para la división
•• para la potenciación, o sea, elevar a una potencia
( ) paréntesis para operaciones preferenciales tal como en álgebra.

Un programa de LP/1 para la raíz de una ecuación cuadrática es:

```
TOME LOS DATOS (A, B, C);
RAÍZ=(—B)+RACU(B** 2—4* A*
C)/(2* A);
COLOQUE DATO (RAÍZ);
```

El punto y coma separa los planteos en programas de LP/1. TOME es un verbo para la entrada. Puede traducirse por "TOME los valores de las variables entre paréntesis". DATOS es una forma particular del planteo TOME que describe cómo deben prepararse los datos para la entrada. Esa forma es:

A = valor B = valor C = valor;

o

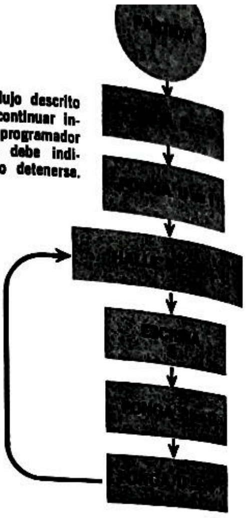
A = 1 B = 4 C = 2;

para los valores 1, 4, y 2 respectivamente.

COLOQUE es el verbo para la salida. DATOS

En esta figura de la animación de una escena de tenis tomada de una computadora de TV, la imagen ha sido trasformada en un despliegue de colores casi abstractos.

Este diagrama de flujo descrito en el texto podría continuar indefinidamente. El programador de la computadora debe indicarle a ésta cuándo detenerse.



es una forma particular de esa salida, y es:

RAÍZ = valor;

o

RAÍZ = 3,414

con los datos de este ejemplo.

El programa completo lee valores para las variables, computa la raíz e imprime su valor. El LP/1 no es directamente ejecutable por la UCP, pero se convierte en instrucciones en lenguaje de máquina por un programa denominado *compilador*. El programa del compilador toma un solo planteo, como, por ejemplo, el de RAÍZ y lo compila en muchas instrucciones en lenguaje de máquina. La escritura en LP/1 está muy alejada de las características de la computadora y más aproximada al lenguaje del campo del problema: en este caso, álgebra.

CORRECCIÓN DEL PROGRAMA

Aun después que hayamos aprendido lenguaje de programación cometeremos errores al escribir un programa, y es menester hallarlos y corregirlos. Durante la compilación del programa por la computadora, el compilador tiene cierta posibilidad de detectar errores e imprimir sugerencias útiles acerca de éstos en forma de mensajes denominados "diagnósticos".

Por ejemplo, podemos escribir descuidadamente el enunciado para computar RAÍZ:

```
RAÍZ=(—B + RACU (B** 2—4* A*
C)/(2* A);
```

Un mensaje de diagnóstico dirá algo acerca del paréntesis faltante. La función RACU tiene su argumento encerrado entre paréntesis como

RACU (B** 2—4* A* C)

y esto deja el paréntesis de la izquierda antes de —B sin su correspondiente paréntesis de la derecha, un error que el compilador puede detectar y lo hará.

Podemos cometer errores más sutiles. Por ejemplo, escribir el enunciado como

```
RAÍZ=(—B + RACU(B** 2—4* A*
C)/2* A;
```

El compilador aceptará esto como un enunciado correctamente escrito, y la computadora lo ejecutará en la forma que lo está. Sin embargo, obtendremos una respuesta errónea. Ahora, la computadora no da sugerencia alguna acerca de por qué el resultado es erróneo. La corrección se vuelve más difícil.

Recuérdese que la computadora seguirá las direcciones implícitas en la forma en que se escribe la ecuación. Computará el numerador de la ecuación correctamente. Luego dividirá el numerador por 2 y luego multiplicará el cociente por A. Conforme está escrito, el enunciado de programación realizará lo siguiente:

$$\text{RAÍZ} = \frac{-B + \sqrt{B^2 - 4AC}}{2} \cdot A$$

La razón es simplemente que el producto 2* A no está encerrado entre paréntesis. Al hacer esto se indica que el numerador debe dividirse por el producto, que es el denominador de la ecuación. A menos que incluyamos 2* A entre paréntesis, no habremos escrito correctamente la instrucción a la computadora. La computadora procesará lo que hemos ideado y programado en ella. Lo cual, en este caso, se entiende en álgebra, pero no en computación.

Podemos cometer otro tipo de error. Supongamos que los valores de A, B y C son:

A = 3 B = 5 C = 3;

para el programa de LP/1. La computadora, en muchos casos, imprimirá un mensaje de diagnóstico cuando trate de computar la raíz cuadrada, porque el valor del radical $B^2 - 4AC$ es —11, o sea, un número negativo. Pero no existe la raíz cuadrada de un número negativo entre los números reales.

Como programadores, debemos tener en cuenta la posibilidad de un valor negativo. Debemos dirigir la computadora para computar el valor como una cantidad aislada y luego comprobar si es negativa antes de seguir adelante con el cálculo de la raíz cuadrada. La responsabilidad de esto cae sobre el programador. De hecho, las instrucciones para computar la raíz nuevamente en la sección anterior sobre algoritmos deberían haber incluido un planteo acerca de qué hacer cuando se computa un valor negativo, pero no lo incluyeron.

Un programa corregido muestra una forma en que esto puede manejarse. La elipse (...) indica que se eliminan algunos otros planteos en el programa.

```
TOME LOS DATOS (A, B, C);
CARACTERÍSTICA =
   (B** 2—4 * A * C);
SI LA CARACTERÍSTICA < 0
   ENTONCES VAYA A ERROR;
   O SI NO, RAÍZ =
   (—B + RACU(CARACTERÍSTICA))/
   (2 * A);
COLOQUE LOS DATOS (RAÍZ); ...
   ERROR: COLOQUE LOS DATOS (CA-
   RACTERÍSTICA); ...
```

Esta versión examina el valor de la característica, y si es negativa salta por encima del cálculo de la raíz a otra parte del programa que imprime el valor de la CARACTERÍSTICA. Si el valor bajo el radical no es menor de cero, el programa procede a computar la RAÍZ.

Un programa en FORTRAN se desarrollaría siguiendo líneas similares, pero no idénticas. Ahora vayamos a un problema diferente para ilustrar un desarrollo más artificioso de un algoritmo.

ALGORITMO PARA LA SERIE DE FIBONACCI

Una serie interesante de números que aparece en muchos libros de álgebra como en libros de introducción a la ciencia de la computación es la sucesión de Fibonacci:

0, 1, 1, 2, 3, 5, 8, 13, 21, 34, ... En esta sucesión de números, se dan 0 y 1 como los dos primeros términos. Luego, cada término se obtiene sumando los dos precedentes.

Se puede continuar generando esta serie tanto como se quiera según la explicación recién dada. Para escribir un algoritmo de computación para este proceso debemos ser mucho

más explícitos. Un algoritmo computacional puede expresarse por medio de un diagrama de flujo, que es un recurso gráfico para representar el procedimiento por pasos que será seguido por la computadora.

Al computar los términos de esta serie se notará inmediatamente que sólo se usan dos para generar el siguiente. Ésos son el "último término" y el "penúltimo término". Después de calcular cada término, el "último término" se vuelve el "penúltimo término", y la suma se vuelve el "último término". Esto se ve fácilmente deslizando un dedo a lo largo de la fila de términos a medida que la serie se desarrolla. Ésta es también la parte artificiosa del planteo del algoritmo computacional.

Adoptaremos la siguiente notación, o símbolos, para los términos:

P para el "penúltimo término"

U para el "último término"

S para la suma.

En la página 480 se muestra un diagrama de flujo literal. Las flechas indican la dirección que llevan los pasos. La flecha que vuelve al bloque con "Halle S" forma un *ciclo* en el programa.

Después de cada computación de la S, el anterior último término U se convierte en el valor del penúltimo término P. La suma S se convierte entonces en el valor del último término U. Luego, todo el proceso recomienza con estos nuevos valores para P y U. Se computa una nueva S. El mismo conjunto de instrucciones se repite una y otra vez en un ciclo.

El diagrama de flujo mostrado en la página 480 insinúa una pregunta. ¿Cómo o cuándo se detiene el proceso? Puede continuar indefinidamente. Debe indicársele a la computadora que se detenga, y esto corresponde al programador. Podemos detenerla después de computar un determinado número de términos de la serie, o cuando el nuevo término alcanza determinado valor. Se necesita algún recurso o regla para detener la computación cíclica.

Los diagramas de flujo para la programación de computadoras son más formales que el literal recién mostrado. Las formas de los bloques llevan implícito un significado cerca de la operación que contienen. Una pequeña flecha indica que un valor debe asignársele a otro; por ejemplo,

$$N \leftarrow 0$$

significa asignar el valor 0 a la N. Un bloque en forma de rombo indica el examen de una condición. Un rectángulo con la base curva designa una salida impresa.

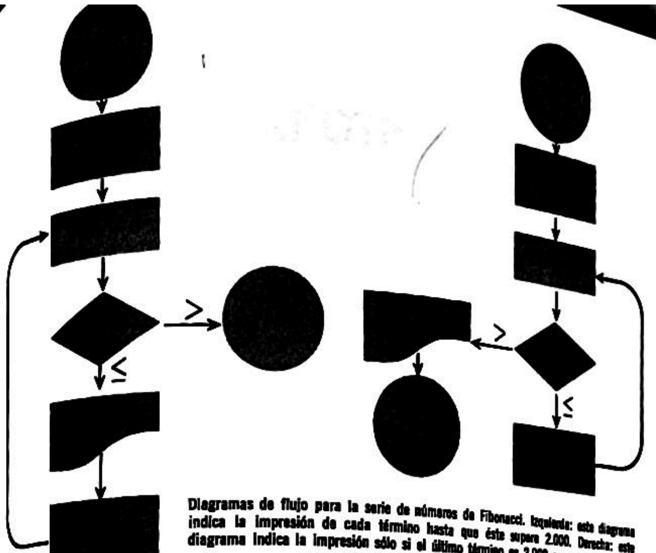
Mostramos, en la página opuesta, dos diagramas de flujo para la serie de Fibonacci. El primero exige la impresión de cada término hasta que éste exceda un valor de 2.000. El segundo exige una impresión sólo si el último término es 2.000 o menos. Cada uno es un algoritmo computacional válido para la serie.

Un programa de computadora puede escribirse directamente del diagrama de flujo en cualquiera de los distintos lenguajes de máquina. El proceso que debe seguir la computadora está completamente expresado en el algoritmo. El programa es simplemente una traducción del algoritmo a algún lenguaje que pueda finalmente ejecutar la computadora.

MUCHAS ACTIVIDADES DE PROGRAMACIÓN

Dentro de la parte que compete al programador de computación –el software como oposición al hardware– han surgido varios niveles de actividad. En el nivel más bajo hay un codificador que es bien versado en un lenguaje de programación y que toma los algoritmos ideados por algún otro y los convierte en un programa de computadora. Más arriba en la escala de actividades están aquellos que trabajan a nivel de algoritmo. Ellos se ocupan de cómo organizar la solución. También pueden hacer alguna codificación, pero esto no es un requisito. Más arriba aún están aquellos que deciden qué problemas resolver. Para este tipo de trabajo se necesita cierto conocimiento y competencia en artes de programación. Más importante es la educación para ayudar a reconocer qué problemas necesitan solución y para comprender cómo puede desempeñar la computadora su papel en esta solución.

Más y más escuelas superiores y colegios están ofreciendo instrucción en la ciencia de computación que incluye programación. Se dispone de muchos libros de programación, incluyendo materiales de instrucción. Se han desarrollado cursos que usan la computadora para enseñar programación. Este campo a veces se denomina IGC, que significa Instrucción Guiada por Computadora. Las lecciones se presentan en un terminal de computadora, que puede ser tan simple como una máquina de escribir, o



Diagramas de flujo para la serie de números de Fibonacci. Izquierda: este diagrama indica la impresión de cada término hasta que éste supera 2.000. Derecha: este diagrama indica la impresión sólo si el último término es 2.000 o menos.

tan complejo como un instituto de aprendizaje con una máquina de escribir, pantalla de TV, proyector de diapositivas y grabador de cintas. Las respuestas a los materiales que se presentan constituyen un diálogo con la computadora, o sea, con un programa de computadora. Las respuestas se evalúan como correctas o incorrectas. Deben darse las instrucciones para releer una sección anterior o para detenerse, y pensar antes de contestar a la misma pregunta otra vez. Esto depende de la estrategia de instrucción que se programa en la computadora.

TIEMPO COMPARTIDO

La programación de lenguajes adecuados a la operación de terminales y la enseñanza se organizan en forma diferente de aquella en que se propone un programa en su totalidad. El nombre de este modo de operación del sistema de computadora es tiempo compartido. Un número de usuarios, cada uno en su propio terminal, está usando una computadora central simultáneamente. En sistemas de tiempo compartido se utilizan lenguajes interactuantes. Puede ser una interacción entre FORTRAN o LP/1, un BASIC o un JOSS.

La cualidad esencial del lenguaje interactuante es que tan pronto como se introduce en la computadora un planteo para usar en un programa, este planteo se evalúa para detectar errores. Los diagnósticos de error aparecen inmediatamente. El planteo puede corregirse antes de continuar con el siguiente.

LA PROGRAMACIÓN ES NECESARIA

La programación es necesaria para el trabajo de computación porque la computadora sólo llevará a cabo un proceso que ha sido expresado en forma de lenguaje de programación. La computadora estará ociosa hasta que se la programe para realizar algo. Sólo entonces puede hacer cálculos relámpago, tomar decisiones en fracciones de segundo y descender hombres en la Luna.

PROCESAMIENTO DE PALABRAS

Las procesadoras de palabras aceleran electrónicamente la escritura, la revisión y la trasmisión de documentos al no tener que volver a leer todo el texto que se ha revisado previamente. Es lógico que ocurran mucho menos errores si sólo las partes donde hubo alguna equivocación requieran que se vuelvan a escribir. Esta ayuda no es nada más para los escritores y los editores, sino también para las secretarias y para aquellas personas que tienen necesidad de utilizar la palabra escrita como medio de comunicación.

La revolución de las procesadoras de palabras comenzó en el año de 1964 con la aparición en el mercado de la máquina de escribir *Selectric*, de cinta magnética y memoria, construida por la empresa IBM. Esta procesadora permitía al usuario grabar y reproducir documentos tantas veces como fuera necesario. Con esta máquina y otras similares era también posible revisar el texto y hacer las correcciones pertinentes. En 1969 apareció la máquina de escribir *Selectric* con tarjeta magnética. Con ella, los documentos se imprimían en tarjetas que podían almacenar hasta una página completa de texto. Si había algún error, éste se tachaba y el cambio quedaba almacenado en la tarjeta. Una vez que se había revisado totalmente el texto se colocaba en la máquina de escribir una hoja de papel en blanco, se oprimía un botón y se obtenían de manera automática copias impecables. Dicha máquina reproducía el texto a razón de 150 palabras por minuto.

Hoy día, otros sistemas de impresión electrónica que se valen de tarjetas, cintas magnéticas, cassettes o mecanismos que pueden manejar la información, ya sea de tarjetas o de cintas, son elementos importantes del mercado de procesamiento de palabras. Otros sistemas más sofisticados pueden poner en orden alfabético los documentos y servir de "archiveros electrónicos". Otras, equipadas con mecanismos que imprimen en ambas direcciones sirven como aparatos de duplicación y de impresión de alta velocidad. Pero no para aquí la cosa. Hay otros aparatos que pueden insertar información variable en formatos estándar, como en una carta. Características adicionales hacen de las procesadoras de palabras ayudantes valiosísimos. Los rastreadores, por ejemplo, permiten al operador buscar, localizar y cambiar rápidamente en un documento una palabra o una frase dada. Tanto los márgenes como los tabuladores, el espaciamiento entre líneas, la alineación de columnas, el subrayado de palabras y el justificado del texto en el margen derecho se pueden obtener de manera automática en estas máquinas.

CON PANTALLA . . .

Hay procesadoras de palabras con ventanillas para permitir la revisión del texto línea por línea. También hay unidades que cuentan con video, es decir con pantalla de rayos catódicos, tipo televisión, en donde se puede leer una página completa y que recorren el texto para su lectura y revisión. Los sistemas globales de localización permiten que una palabra, o una frase específica, se pueda cambiar cada vez que aparece en un documento. Por ejemplo, se puede hacer que el procesador cambie la dirección del señor Pérez, de la calle del Olmo 37 a la de Margaritas 425 cada vez que aquélla se repita. Es más, aún párrafos enteros se pueden cambiar de lugar en dicha procesadora.

Otra procesadora de palabras, la *Redactor II*, muestra en su pantalla las secciones corregidas del texto en colores distintos. Si, por ejemplo, el texto se muestra en letras verdes en fondo negro, la sección corregida aparecerá con letras negras en fondo verde.

Los sistemas de video permiten también que el usuario mecanografíe o corrija un documento mientras la máquina imprime manuscritos ya revisados y corregidos previamente. El mecanografiado se efectúa, por lo regular, con la ayuda de una "margarita", es decir el disco impresor que tiene una letra en cada pétalo. Este disco es bidireccional y muy rápido, pues por lo regular escribe unas 360 palabras por minuto. Ahorra tiempo, ya que no tiene que regresar al margen izquierdo cada vez que termina de escribir una línea.

Existen otros sistemas con video que se pueden programar para que procesen automáticamente aquellas frases que se utilizan con frecuencia en la correspondencia. Y hay otras procesadoras con mecanismos de seguridad incorporados en los que se programa una clave para restringir el acceso a toda aquella información clasificada.

. . . Y HABLAN . . .

Algunas empresas están diseñando sistemas de comunicación que se pueden emplear con la procesadora de palabras. Hay unidades con video y máquina de escribir que pueden comunicar por teléfono o por Telex con otras unidades que cuentan con video, con computadora o con aparatos impresores.

. . . Y LEEN . . .

Conforme las procesadoras de palabras se hacen más comunes y más sofisticadas, la distinción entre ellas y las procesadoras de datos es más difícil de definir. El registro de datos y el correo electrónico se han incorporado en algunos sistemas de procesamiento de palabras, en tanto que otras pueden seleccionar información de la misma manera que una computadora.

Junto con la tendencia hacia una mayor sofisticación está la aparición de las máquinas de reconocimiento óptico de caracteres en los sistemas procesadores de palabras. Este reconocimiento óptico fue diseñado originalmente para acelerar el registro de los datos en la computadora. Hoy día hay modelos más pequeños de estas máquinas a la disposición del público para leer el material mecanografiado hasta cuatro páginas por minuto. No obstante, hay inconvenientes: como las máquinas de reconocimiento óptico nada más pueden leer determinados caracteres de imprenta, sólo se pueden utilizar máquinas de escribir que cuenten con dichos caracteres.

. . . Y TRADUCEN . . .

Por último, hay una computadora que ayudará a los traductores a acelerar la traducción de un idioma a otro. Según el fabricante de dicha computadora el traductor podrá manejar más de 2,400 palabras por hora. La computadora hace traducciones literales y el traductor actúa como corrector e intérprete. La máquina en sí es una procesadora de palabras con máquina de escribir electrónica y pantalla conectada a una minicomputadora.

El texto en el idioma extranjero por traducir se alimenta en la máquina y la traducción, palabra por palabra, aparece en la pantalla. Los cambios en la traducción —que serán hechos por el traductor— se mecanografían para corregir la versión procesada. En donde hay más de un significado para una palabra, la computadora da la lista de todas las definiciones para que el traductor escoja la adecuada. El sistema, llamado Procesador de Palabras Multilingües, fue diseñado para traducir del español al inglés, pero ya se cuenta con elementos para traducir el japonés, el ruso, el árabe, el hebreo y otros cinco idiomas al inglés.

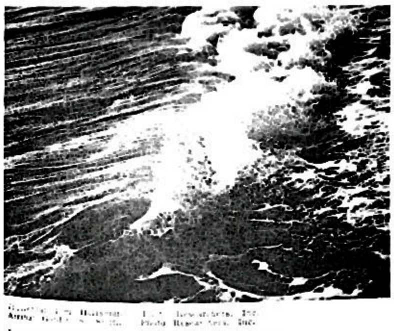
. . . ¿Y OIR? . . . AÚN NO

Con tantos tipos diferentes de procesadoras de palabras en el mercado, el comprador se enfrenta con el problema de encontrar un equipo adecuado que no se haga obsoleto en poco tiempo. Por consiguiente, los fabricantes producen equipo que es compatible con sistemas anticuados, pero que actualizan sistemas modulares, permitiendo así al comprador la misma flexibilidad que existe en los sistemas modulares de procesamiento de datos.

Por último, se espera que pronto haya procesadoras de palabras con voz. Tal sistema permitiría que los documentos que se van a imprimir fueran escuchados al mismo tiempo que se proyectan en la pantalla del aparato. Un sistema de computación comprobaría posteriormente el documento en lo que a ortografía se refiere y, por último, después de archivar electrónicamente la copia, trasmitiría el documento por correo electrónico.

Así, el procesamiento de palabras que se inició con una máquina de escribir ha permitido que el usuario mecanografíe el texto sin volver a copiarlo enteramente, y no está lejano el día en que una máquina parlante facilite rápidamente la revisión y la distribución de los documentos sin necesidad de utilizar una pluma, una máquina de escribir e incluso papel.

CIENCIA DE LA TIERRA



La faz de nuestro planeta está siendo constantemente cambiada por las fuerzas que crean volcanes activos, como el Irazú (izquierda), en América Central, y producen ondas que baten las costas (arriba).

194-201 **El estudio de la Tierra**
202-209 **Corteza terrestre**
210-219 **Deriva continental**
220-229 **Terremotos**
230-238 **Minerales**
239-243 **Coleccionemos minerales y rocas**
244-252 **El interior de la Tierra**
253-266 **Volcanes y géiseres**
267-277 **Las montañas**
278-285 **Llanuras, mesetas y desiertos**
286-296 **Cavernas**

H. W. Kitler, Audubon PR

Foto de la Fuerza Aérea de EU

EL ESTUDIO DE LA TIERRA

Antes del siglo XVII se creía que la Tierra era el centro de todas las cosas, el eje en torno del cual giraban Sol, Luna, estrellas y planetas. Desde entonces, esta teoría geocéntrica –o "centrada en la Tierra"– se dejó de lado. Ahora sabemos que nuestro mundo es un pequeño planeta que, junto con sus hermanos, describe una órbita en torno de una estrella llamada Sol, de la que también sabemos que es una entre miles de millones de estrellas en este "universo-isla" conocido como la Vía Láctea, que, a su vez, es sólo uno entre miles de millones de universos-islas. Evidentemente, nuestro planeta es sólo un punto insignificante en el Universo, y no su centro; pero, diminuto como es, tiene suprema importancia para nosotros, por ser nuestro hogar en el espacio.

EL ORIGEN DE LA TIERRA

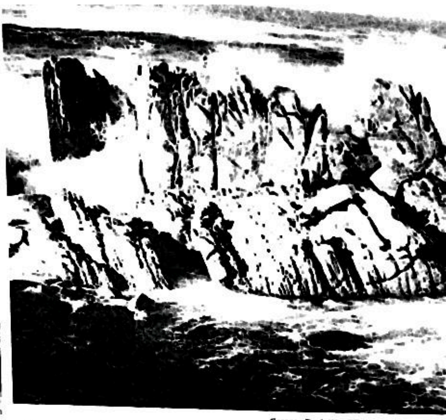
¿Cómo se originó la Tierra? Para dar una respuesta satisfactoria a esta pregunta se han realizado muchos intentos desde la más remota antigüedad, algunos muy ingeniosos. Sin embargo, hasta el siglo XVIII no hubo una hipótesis suficientemente válida para que los modernos hombres de ciencia la consideraran digna de discusión.

En 1755 el gran filósofo alemán Immanuel Kant sugirió que el Sistema Solar (Sol, planetas, satélites, cometas y demás cuerpos celestes) se formó a partir de una nebulosa, esto es, una gran masa de gas tenue y vaporoso. La teoría de Kant no causó mucho revuelo en el mundo científico.

Por la misma época, el naturalista francés George-Louis Leclerc, conde de Buffon, dio su opinión personal sobre cómo nació la Tierra. Según él, muchísimo tiempo atrás, el Sol chocó con un cometa y, como resultado, una gran cantidad de material fue arrancado de él; después, este material se enfrió y dio origen a los planetas. Infortunada fue la elección de un cometa hecha por Buffon, ya que estos astros son muy insignificantes en comparación con el Sol, al cual tal encuentro no lo habría afectado. No obstante, la teoría de la colisión que propuso habría de servir como prototipo para varias hipótesis modernas basadas en la idea del choque entre cuerpos celestes.

Hipótesis nebular. Pierre Simon, marqués de Laplace, astrónomo matemático francés, rechazó la teoría de Buffon y, en 1796, ofreció una propia. Se la llamó hipótesis nebular y se la aceptó ampliamente hasta fines del siglo XIX. Recuerda la hipótesis nebular de Kant en varios detalles, pero Laplace probablemente no conocía la contribución de aquél.

Exxon



George Daniell, PR

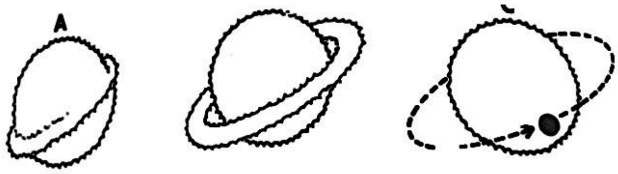
El estudio de la Tierra incluye el análisis de los procesos naturales que han trasformado y continúan trasformando la superficie. En el proceso del vulcanismo, materia rocosa fundida, o magma, proveniente del interior del planeta, se trasforma en rocas. A veces, magma, vapor de agua y otros gases escapan de un volcán (página de enfrente, izquierda). En la agradación, se generan depósitos cuando partículas rocosas son arrastradas por los glaciares y otros agentes (página de enfrente, derecha). En el diastrofismo se desplazan partes de la corteza, causando muchos cambios, incluso plegamientos en las capas de roca (arriba, a la izquierda). En la degradación, las rocas son desgastadas por diversas fuerzas, entre ellas el embate de las olas (arriba, a la derecha).

Según Laplace, en un tiempo los miembros de la familia solar formaron parte de una enorme masa de gases incandescentes en lenta rotación, la que gradualmente se enfrió y contrajo, adquiriendo una forma cada vez más esférica. A medida que rotaba con rapidez creciente se fue abultando alrededor del ecuador. Finalmente, desde esta zona fue despedido un anillo de materia que se enfrió y se contrajo, trasformándose, por último, en un planeta cuya órbita estaba en el plano que primitivamente ocupaba el anillo. Desde la masa central fueron expelidos anillos uno tras otro que evolucionaron y, con el tiempo, se trasformaron en planetas; la masa central se convirtió en nuestro Sol, y a su vez los planetas lanzaron anillos al espacio, los que se trasformaron en satélites planetarios, o lunas. Por tanto, según la hipótesis nebular, el origen de la Tierra habría sido un anillo de materia lanzado desde una masa central que luego se convirtió en el Sol.

Hipótesis planetesimal. Alrededor de 1900, el astrónomo Forest Ray Moulton y el geólogo T. C. Chamberlin, ambos de la Universidad de Chicago, presentaron una nueva teoría, a la que llamaron hipótesis planetesimal. Un planetésimo es un pequeño cuerpo sólido que gira en torno de un núcleo gaseoso. Según Moulton y Chamberlin, una estrella que se desplazaba velozmente por el espacio se acercó mucho a nuestro Sol. El gran aumento de las fuerzas gravitatorias entre ambas estrellas hizo que cada una levantase enormes mareas en el cuerpo caliente y gaseoso de la otra.

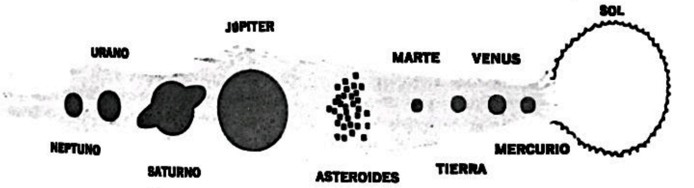
A medida que las mareas solares ocasionadas por la atracción de la estrella pasajera se hacían más grandes, desde el Sol fueron expelidas masas de gas que comenzaron a girar en sucesivas vueltas; algunas siguieron a la otra estrella en su rápida carrera hacia el espacio; otras, retenidas por la atracción del Sol, empezaron a girar en torno de él. Cuando la otra estrella se alejó, se apaciguaron las grandes mareas solares. Las masas de gas arrancadas del Sol se asentaron en trayectorias ordenadas en torno de él. A medida que se volvieron más frías, se licuaron y, gradualmente, se convirtieron en pequeñas masas sólidas. Estos fragmentos, o planetésimos, con el tiempo se reunieron para formar planetas. Según esta teoría, los meteoritos de nuestro Sistema Solar representan planetésimos que no participaron en el proceso planetogénico.

Teoría de las mareas. En 1918, sir James Jeans y H. Jeffreys, científicos ingleses, elaboraron la teoría llamada de las mareas. También se basaba en la idea de una colisión. Estos científicos diferían de Moulton y Cham-



La formación de los planetas, según la hipótesis nebular de Laplace. Una gran masa de gas incandescente en rotación se enfrió gradualmente y se contrajo, volviéndose cada vez más esférica. A medida que rotaba más rápido, desarrolló un abultamiento a lo largo del ecuador (A). Finalmente, se desprendió un anillo de materia (B). Esta masa se enfrió, se contrajo y se convirtió en un planeta sólido, con su órbita en el plano originalmente ocupado por el anillo.

Según la teoría planetesimal de Moulton-Chamberlin, una estrella se aproximó al Sol y atrajo nubes de gas de él, las que se enfriaron gradualmente hasta convertirse en pequeñas masas sólidas: los planetésimos, que luego se unieron para formar planetas.



Teoría de la marea de Jeans y Jeffreys, quienes sostenían que los planetas se formaron a partir de un filamento de gas, en forma de cigarro, arrancado del Sol por una estrella pasajera. Los planetas más pequeños derivaron de los extremos del filamento; los más grandes, de la sección central. Como Plutón no se conocía cuando se propuso esta teoría, no se lo incluye en la ilustración.

Lyttleton sostuvo que el Sol fue antes una estrella doble (A); y que otra estrella, al pasar, desintegró una del par, trasformándola en una masa de gas arremolinado (B), que luego evolucionó, originando los planetas.

berlin en que no creían que los planetas derivasen de grandes cantidades de cuerpos pequeños, o planetésimos; opinaban que los planetas se formaron directamente a partir de la masa original de gas atraída desde el Sol por una estrella pasajera, y no por generación de grandes cuerpos sólidos a partir de pequeñas partículas. Según la teoría de las mareas, cuando la estrella se aproximó o, inclusive, rozó tangencialmente a nuestro Sol, su atracción gravitatoria arrancó de él un largo filamento de gas en forma de cigarro; esto es, más ancho en la sección media que hacia los extremos (ver diagrama). Los planetas más pequeños, es decir, los más próximos, y los más alejados se formaron a partir de los extremos del filamento; los más grandes, de su parte media.

Teoría de Lyttleton. El astrónomo R. A. Lyttleton ofreció aún otra modificación de la teoría de la colisión: sugirió que, originalmente, el Sol era una estrella doble, es decir, un sistema formado por dos astros que giraban en torno de un centro común de gravedad. Una estrella que pasó pudo haberse acercado a uno de los dos soles, desintegrándolo y trasformándolo en una vasta extensión de gases remolinantes. La estrella que sobrevivió sería nuestro Sol. La víctima de la colisión se habría gradualmente convertido en nuestros planetas. En cierto sentido, la hipótesis de Lyttleton explica mejor el origen del Sistema Solar que las otras versiones de la teoría de la colisión.

Muchos astrónomos modernos se inclinan a descartar las teorías basadas en una colisión, o semicolisión, entre el Sol y una estrella viajera; creen que el Universo, como un todo, ha evolucionado de modo gradual y ordenado, y no a través de catástrofes accidentales extrañas al curso normal de los acontecimientos. La antigua hipótesis nebular de Kant y Laplace, convenientemente modificada, ofrecería, justamente, tal proceso ordenado y uniforme.

Modificaciones de la hipótesis nebular. El astrónomo alemán C. von Weizsaecker presentó una hipótesis nebular propia en la década de 1940. Sostenía que, al principio, desde el ecuador del Sol primigenio se extendió un velo de materia gaseosa hacia el exterior, hasta grandes distancias. La mayor parte de este velo estaba constituida por los elementos livianos hidrógeno y helio. Con el tiempo, el calor y la presión de la radiación solar expulsaron gran parte de estos dos elementos, dejando los más pesados que, gradualmente, se reunieron en una serie de masas concéntricas de forma arriñonada; éstas atrajeron otras sustancias del espacio y evolucionaron hasta ser planetas.

El astrónomo neerlandés-norteamericano G. P. Kuiper propuso otra hipótesis nebular. Supuso que, originalmente, hubo una nebulosa de forma discoidal y enorme extensión, con el protosol –el futuro Sol– en el centro. La composición total de la nebulosa era uniforme. La temperatura era baja porque el protosol todavía no había comenzado a irradiar. Esta nebulosa fría comenzó a desmembrarse y concentrarse en masas separadas: los protoplanetas, o futuros planetas. La materia central –el protosol– también se concentró bajo la fuerza gravitatoria; a medida que se contraía, aumentaba su temperatura; su energía radiante expulsó a la mayoría de los elementos más livianos (particularmente hidrógeno y helio) fuera de los protoplanetas y de la propia nebulosa. En cada protoplaneta, la mayor parte de los elementos más pesados, como hierro, níquel y varios otros metales, se acumuló en el centro. A medida que los protoplanetas disminuían su tamaño, comenzaron a rotar cada vez más rápido, alcanzando la mayor parte del momento que tienen en la actualidad.

Teoría de la nube de polvo. El astrónomo norteamericano Fred L. Whipple propuso una teoría del Universo basada en una nube de polvo: según Whipple, el futuro Sistema Solar primero fue una vasta nube de gases y polvo cósmico que tomó forma discoidal. Irregularidades dentro de la nube originaron su rotación. Polvo y gases se concentraron al girar, y la nube se deshizo. Las partículas sólidas que contenía chocaron y se unieron entre sí, trasformándose en planetas. Los gases del centro de la primitiva nube originaron el Sol.

Éstas son sólo unas pocas entre las muchas teorías propuestas para explicar el nacimiento de la Tierra. Ninguna de ellas puede considerarse totalmente satisfactoria, y ninguna está aceptada universalmente en la actualidad.

LA ESTRUCTURA DE LA TIERRA

Nos encontramos en terreno más firme cuando analizamos la estructura y la composición de nuestro planeta, pues los científicos

han podido acumular un impresionante conjunto de pruebas. Poseemos información particularmente exacta sobre la relativamente delgada capa exterior de la Tierra, o corteza, porque parte de ella se ha examinado en forma directa.

LA CORTEZA

La corteza terrestre está constituida por rocas formadas de diversos minerales. Existen muchas clases diferentes de ellas, pero se las puede dividir en tres tipos básicos: ígneas, sedimentarias y metamórficas.

Las rocas *ígneas* (literalmente, "rocas producidas por el fuego") fueron antes sustancias fundidas sumamente calientes que se solidificaron ya sea en la superficie de la Tierra, como el basalto, o muy profundamente dentro de ella, como el granito. Más adelante volveremos a hablar de las rocas ígneas, que constituyeron la corteza terrestre original.

Las rocas *sedimentarias* derivan de fragmentos rocosos desplazados y diseminados por el viento, el agua u otros agentes. Depositadas en tierra o en el lecho marino, gradualmente fueron unidas por presión y aglutinadas. Las rocas sedimentarias forman capas que van desde unos pocos centímetros hasta varios metros. Constituyen la mayor parte de la superficie de la corteza terrestre. Entre las rocas más importantes están la arenisca, la caliza y la pizarra.

Ciertas capas que se encuentran en las rocas sedimentarias han derivado de los restos de plantas que una vez prosperaron en zonas pantanosas. Los restos, sepultados bajo sedimentos posteriores, fueron sometidos a elevado calor y gran presión; finalmente, se convirtieron en carbón. Los animales han ayudado a crear varias clases de rocas sedimentarias: la caliza –al menos en parte– ha derivado del carbonato de calcio proveniente de las valvas de caracoles de mar, almejas, ostras y otros animales marinos. Tanto las plantas como los animales han contribuido a la formación de arrecifes e islas de coral. Otras formaciones sedimentarias fueron producto de la evaporación del agua; los compuestos contenidos en solución en el agua del mar constituyeron depósitos a medida que se producía la evaporación. Por ejemplo, la sal gema y otras sales solubles derivadas de la evaporación se encuentran ahora en rocas sedimentarias de distintas partes del mundo; se presentan en lechos cuyo espesor total va desde un metro hasta más de 300.

Los restos e improntas vegetales y animales –los llamados *fósiles*– contenidos en la roca sedimentaria revelan mucho sobre su formación. Entre otras cosas, indican si los sedimentos fueron depositados en tierra, sobre los fondos oceánicos, o de otras masas de agua. Por lo común, los acumulados en tierra contienen fósiles de animales y plantas terrícolas; los que están en el mar, valvas de animales marinos. Asimismo, los fósiles dan datos sobre el clima en el momento de la deposición. Evidentemente, si en una roca sedimentaria hay restos de animales o plantas tropicales, ello indica que se depositaron cuando el clima era cálido. Los fósiles también ofrecen información invalorable sobre la historia de la flora y fauna terrestres.

Las rocas de la tercera división principal, o *metamórficas*, están constituidas por rocas ígneas o sedimentarias que se han trasformado en el trascurso de los tiempos. En el metamorfismo, o formación de estas rocas, entran en juego varios factores: presión, calor, presencia de agua y alteraciones químicas. Las partículas de roca original son forzadas a ordenarse en nuevas formas. Pueden dar nuevos minerales; a veces, la roca metamórfica retiene trazas de la formación rocosa de la que deriva, a veces es una roca completamente nueva. Entre las rocas metamórficas están el mármol, la pizarra y los gneis graníticos. El mármol deriva de la caliza; la pizarra, de la lutita; los gneis graníticos, de diversos granitos.

En las zonas de tierra firme, la capa superior de la corteza está generalmente cubierta por tierra vegetal (humus), constituida en gran parte por partículas de rocas desgastadas. La tierra vegetal también contiene considerable proporción de materia orgánica, proveniente de los productos en descomposición de seres de épocas anteriores. La tierra vegetal mantiene la flora y también la fauna terrestres, pues el alimento de los animales proviene, directa o indirectamente, de las plantas.

INTERIOR DE LA TIERRA

La corteza terrestre se extiende hacia abajo relativamente pocos kilómetros: alrededor de 30 a 40 en las zonas continentales, y sólo de 5 a 10 bajo los principales lechos oceánicos. Debajo de la corteza se encuentra el más o menos misterioso interior de la Tierra. Como hasta ahora no hemos podido obtener información directa sobre esta región, debemos basarnos en pruebas indirectas para estudiarla. Debemos analizar, por ejemplo, los registros obtenidos de ondas sísmicas que pasan a través del interior de la Tierra y que son afectadas de diversos modos por las distintas capas a través de las cuales se desplazan.

El interior de nuestro planeta consiste en dos regiones principales: un núcleo central de níquel y hierro, con un diámetro de unos 7.000 kilómetros, y un manto circundante, con un espesor de cerca de 3.000. Los científicos han calculado la densidad del interior de la Tierra en varios niveles, su temperatura y elasticidad; han hecho un análisis de la composición de los materiales contenidos en estas regiones; también han teorizado sobre la generación del campo magnético del planeta, muy por debajo de la corteza.

TRES ESFERAS DE LA TIERRA

Hasta aquí hemos discutido la estructura rocosa básica de la Tierra. En realidad, las rocas (y el suelo, que proviene en gran medida de aquéllas) sólo forman una de las tres provincias, o "esferas", en las que se divide nuestro planeta. Rocas y suelo son la parte sólida: litosfera ("esfera de roca"). Las aguas, que se encuentran en océanos, mares, lagos interiores, arroyos, aguas subterráneas y formaciones de hielo continental, constituyen la hidrosfera ("esfera de agua"). Finalmente, el océano invisible de aire que rodea la Tierra es la atmósfera ("esfera de aire").

A veces, estas divisiones principales se superponen; así, la atmósfera contiene agua en forma de vapor y partículas rocosas, llamadas polvo. Buena parte del aire es absorbido por la hidrosfera, y gran cantidad de partículas de roca queda en suspensión en sus aguas. La litosfera contiene tanto agua como aire.

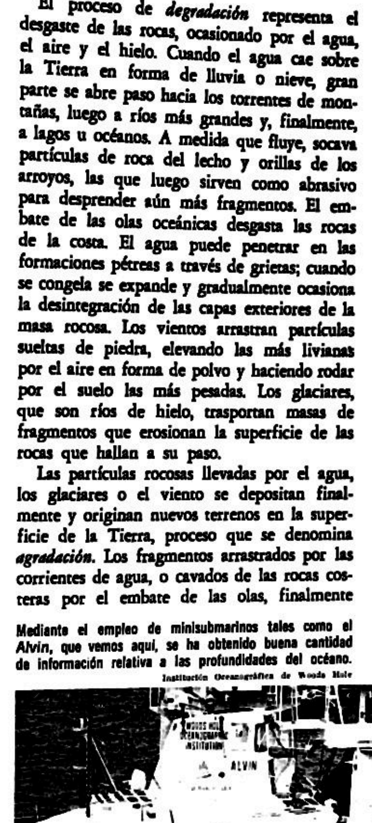
RASGOS CAMBIANTES DE LA TIERRA

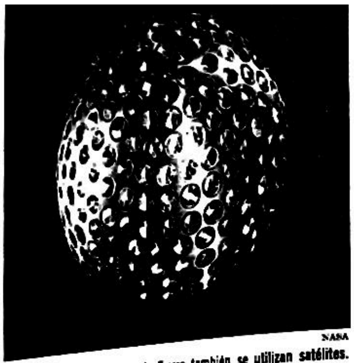
Los que moramos en la Tierra estamos particularmente familiarizados con los rasgos del paisaje de la superficie: llanuras, montañas, océanos, arroyos, cascadas, cañones, puentes naturales y otros elementos. Estos rasgos de ningún modo son estables; los modifican constantemente ciertos procesos naturales que han estado obrando sobre ellos desde los albores de la historia del planeta.

El proceso de *degradación* representa el desgaste de las rocas, ocasionado por el agua, el aire y el hielo. Cuando el agua cae sobre la Tierra en forma de lluvia o nieve, gran parte se abre paso hacia los torrentes de montañas, luego a ríos más grandes y, finalmente, a lagos u océanos. A medida que fluye, socava partículas de roca del lecho y orillas de los arroyos, las que luego sirven como abrasivo para desprender aún más fragmentos. El embate de las olas oceánicas desgasta las rocas de la costa. El agua puede penetrar en las formaciones pétreas a través de grietas; cuando se congela se expande y gradualmente ocasiona la desintegración de las capas exteriores de la masa rocosa. Los vientos arrastran partículas sueltas de piedra, elevando las más livianas por el aire en forma de polvo y haciendo rodar por el suelo las más pesadas. Los glaciares, que son ríos de hielo, trasportan masas de fragmentos que erosionan la superficie de las rocas que hallan a su paso.

Las partículas rocosas llevadas por el agua, los glaciares o el viento se depositan finalmente y originan nuevos terrenos en la superficie de la Tierra, proceso que se denomina *agradación*. Los fragmentos arrastrados por las corrientes de agua, o cavados de las rocas costeras por el embate de las olas, finalmente

Mediante el empleo de minisubmarinos tales como el *Alvin*, que vemos aquí, se ha obtenido buena cantidad de información relativa a las profundidades del océano.
Institución Oceanográfica de Woods Hole





NASA

Para el estudio de la Tierra también se utilizan satélites. Este satélite Geodinámico Laser estudia la rotación terrestre y los movimientos de la corteza.

caen al fondo de ríos, lagos o mares, para formar lodo o arena. Los fragmentos trasportados por glaciares quedan como depósitos cuando se funde el hielo; los llevados por los vientos acaban como dunas de arena o capas de polvo en la superficie. Con el trascurso del tiempo, las partículas trasportadas por el viento pueden llegar a formar capas gruesas. Éste es el origen de la formación llamada loess, que se encuentra en China, Argentina, partes de Europa, el valle del Misisipí en los Estados Unidos y otros lugares.

Ciertos grandes cambios en la superficie terrestre han sido causados por el *diastrofismo:* esto es, por mutuos movimientos de partes sólidas de la corteza. Ésta puede ser combada, plegada, elevada o descendida. Algunos movimientos diastróficos son muy lentos y graduales; otros, como los familiares terremotos, son rápidos y violentos, y pueden súbitamente bajar o levantar grandes masas de terreno, o modificar sus posiciones relativas a lo largo de líneas de falla.

Por doquier existen pruebas del diastrofismo: por ejemplo, vemos la actividad de este proceso cuando examinamos las laderas de acantilados donde han quedado descubiertos estratos sucesivos. En un lugar, están moderadamente inclinados; en otro, plegados, formando una serie de crestas y depresiones onduladas. A veces, ha ocurrido una fractura en un estrato y, a menudo, sus costados se han desplazado uno respecto del otro.

En ciertos lugares el suelo se ha elevado notablemente en el trascurso de los tiempos. Sabemos que las rocas sedimentarias que contienen fósiles marinos tienen que haber estado sumergidas en un momento dado; tales capas pueden hallarse ahora con frecuencia muy por encima del nivel del mar en los Himalayas, en las Rocosas y en el territorio de mesetas a través del cual se ha abierto el Gran Cañón del Colorado. En otras partes, la tierra se ha hundido en la misma proporción: por ejemplo, en el delta del río Ganges, en la India, se han hallado huesos de animales terrestres a cien metros por debajo del nivel del mar. Los antiguos romanos erigieron un templo dedicado al dios Serapis en la costa, no muy lejos de Nápoles, Italia: las pocas columnas que han quedado en pie están horadadas por moluscos marinos hasta una altura de casi seis metros por encima del piso del templo; se han encontrado algunas valvas en los agujeros, lo que demuestra que en esa zona el suelo debe haberse hundido en el mar después de la construcción del templo, arrastrándolo consigo. Más tarde el suelo se elevó de nuevo, y el templo quedó otra vez en tierra firme.

Finalmente, existe el proceso del *volcanismo,* también conocido como *vulcanismo:* es el movimiento de materia rocosa fundida –llamada *magma*– y su trasformación en roca. Una vez se pensó que el calor que hacía fundir las rocas en el seno de la Tierra era un resto del período temprano de su historia, cuando todo el planeta se hallaba en estado de fusión; empero, los geólogos modernos se inclinan a creer que el calor interno se debe, al menos en gran parte, a rocas radiactivas. En el proceso de radiactividad natural, ciertos elementos, tales como uranio y torio, se desintegran espontáneamente, dando otros elementos y cediendo grandes cantidades de energía térmica en el proceso. Se sabe que el uranio y el torio se encuentran, por lo general, en muy pequeñas cantidades en todas las rocas. Los geólogos consideran que el calor liberado por los átomos de uranio y torio en desintegración se acumularía en zonas locales separadas, porque las rocas son malas conductoras; con el tiempo, éste llegaría a elevarse lo suficientemente para fundir las rocas circundantes y producir magma.

Algunos movimientos magmáticos están confinados al interior de la Tierra, y las rocas ígneas resultantes también permanecen allí enterradas, a menos que queden al desnudo por explotaciones de canteras, o por los efectos de la erosión sobre la superficie terrestre. En algunos casos el magma, que se desplaza a lo largo de fisuras en la roca, alcanza la superficie y fluye por encima de ella; o bien puede ser arrojado, junto con grandes volúmenes de vapor de agua y otros gases, por las chimeneas de los volcanes. El magma que se desplaza sobre la superficie de la Tierra se llama lava; ésta se enfría, endurece y trasforma en roca ígnea. Siempre que quedan expuestas en la superficie de la Tierra, las formaciones rocosas provenientes de procesos volcánicos están sometidas al proceso de degradación.

CIENCIAS DE LA TIERRA

Éste es, en síntesis, el siempre cambiante planeta en que moramos. Su estudio es, particularmente, dominio de las que denominamos ciencias de la Tierra, de las cuales la primordial es la *geología* (del griego *ge*: "Tierra", y *logos:* "ciencia"). Examina ésta además de los diferentes procesos que originaron las estructuras rocosas y el paisaje terrestre, la actual composición de la Tierra, tanto exterior como interior. La geología también reconstruye la sucesión de cambios pretéritos en la estructura terrestre, lo que conduce a una comprensión no sólo de la historia de la propia Tierra, sino también de las diferentes formas de vida que la habitaron en cualesquiera de las edades pasadas.

La geología posee varias subdivisiones. La ciencia de la *petrología* estudia el carácter y origen de todas las clases de rocas. La *geología estructural* examina las disposiciones estructurales de las rocas, originadas como consecuencia de plegamientos y fracturas. Los minerales que las constituyen son estudiados por la ciencia de la *mineralogía.* La *estratigrafía* se ocupa de la serie de capas rocosas, particularmente de las sedimentarias. La *geomorfología* considera las diversas formas del terreno que han sido esculpidas por agentes superficiales tales como viento, agua corriente y precipitación. La *glaciología* trata de los procesos ligados al hielo y la nieve. La *paleontología* estudia los fósiles animales y vegetales que se hallan en los terrenos y las pruebas del desarrollo gradual de la vida en épocas pretéritas. Los yacimientos rocosos valiosos para el hombre (de metales, carbón, petróleo y otros similares) son el tema de la *geología económica.* La exploración en busca de éstos depósitos, y su explotación, se llama *geología de minas.* Los ingenieros civiles tienen que conocer las cualidades y resistencia de la Tierra y rocas que hallan, y que también emplean en sus construcciones; este estudio es la *geología de ingenieros.* Según algunos expertos pueden incluirse otras ciencias en la geología. La *geofísica* trata de la física de nuestro planeta. La *geoquímica* se ocupa de la composición de las rocas, minerales, suelo, aire y agua de la Tierra. El terreno en sí pertenece a la ciencia de la *pedología.* La *hidrología* abarca todas las masas de agua –lagos, cursos de agua y pantanos– de los bloques continentales.

La *geología marina,* o *submarina* (*oceanografía geológica*), se ocupa del fondo del mar y los litorales. La delineación exacta de los rasgos de la Tierra es el tópico de la *topografía, de la hidrografía* (para masas de agua) y de la *cartografía.*

Campos afines con la geología son: la *fisiografía* (*geografía física*) que trata de los paisajes de diferentes países; la *ecología* y la *biogeografía,* que se ocupan de la vida en relación con su ambiente, la tierra y el mar; la *limnología,* que trata de la aptitud del agua para sostener vida en su seno; la *oceanografía,* ciencia del mar; *meteorología,* estudio del tiempo; *climatología,* estudio de los climas, y *astrogeología* (*astronomía planetaria*), que se ocupa de la naturaleza de los planetas, satélites, asteroides y meteoritos.

CORTEZA TERRESTRE

Cuando vemos cerros rocosos y farallones escarpados que irrumpen abruptamente en un agradable paisaje de bosques verdes y de campos ondulantes, tendemos a considerar intrusas a estas formaciones rocosas. Sin embargo, constituyen la verdadera corteza de la Tierra, tanto en las zonas marítimas como en las terrestres. La capa fértil del suelo, donde se desarrolla la vida animal y vegetal, proviene de la roca. Se trata de una capa sumamente delgada, compuesta principalmente de roca fragmentada por la intemperie y la erosión. En las zonas oceánicas, depósitos de sustancias minerales y restos de plantas y animales forman también una cubierta superficial sobre el lecho de rocas subyacentes del fondo del mar.

La corteza terrestre es relativamente delgada, de unos 30 a 40 kilómetros de espesor debajo de los continentes, y de menos de 10 kilómetros bajo el suelo oceánico. Un rasgo interesante de la corteza oceánica es su edad. Es más reciente que la situada bajo las zonas terrestres. Muchos geólogos consideran que constantemente se agrega nuevo material a esta corteza, el que proviene del interior de la Tierra y emerge a lo largo de grietas, formando costurones oceánicos. Allí, el piso del mar se expande, y empuja lejos del costurón el material más viejo de la corteza.

La corteza pétrea de la Tierra está compuesta por tres tipos de formaciones rocosas: rocas ígneas, sedimentarias y metamórficas. Las ígneas se han formado por solidificación de material pétreo fundido. Las rocas sedimentarias se componen de sedimentos, o fragmentos, convertidos en rocas en el curso de los siglos. Las metamórficas fueron originariamente rocas ígneas o sedimentarias, trasformadas luego por los cambios de temperatura y presión, y por otros factores que actúan dentro de la corteza terrestre.

ROCAS ÍGNEAS

La masa fundida hirviente de la cual se han formado las rocas ígneas se denomina *magma*. Contiene varios gases en solución, pero en su mayor parte está compuesta por minerales disueltos al azar.

De acuerdo con algunas teorías sobre la formación de la Tierra, nuestro planeta fue una vez una turbulenta masa de gases expulsados por el Sol. A medida que esos gases se enfriaban poco a poco, se convertían en un líquido. Este líquido, el magma, comenzó a enfriarse, y los minerales que contenía empezaron a cristalizar. Los minerales pesados tendían a hundirse en la masa todavía líquida del magma. Los más livianos flotaban por encima de los más pesados. A medida que el enfriamiento continuaba, las rocas comenzaban a solidificarse, y con el tiempo formaron una corteza sólida.

Esto fue sólo una etapa en la formación de las rocas ígneas. Grandes cantidades de magma, con temperaturas sumamente altas, quedaron aprisionadas bajo la corteza. Los gases contenidos en el magma ejercieron presiones enormes, no sólo sobre las rocas ígneas ya formadas, sino también sobre un tipo totalmente distinto de material: el sedimentario.

Las rocas ígneas originales y las sedimentarias que se formaron después no pudieron resistir algunas veces la presión del magma situado bajo la corteza. La roca líquida se abrió entonces camino hacia niveles más altos.

En algunos casos, el avance del magma quedó detenido por las capas de roca que encontraba a su paso, y sólo sus gases escaparon. Poco a poco, el magma se enfrió y solidificó bajo las capas de rocas que formaban la parte superior de la corteza terrestre. Las rocas ígneas formadas de esta manera se denominan masas intrusivas.

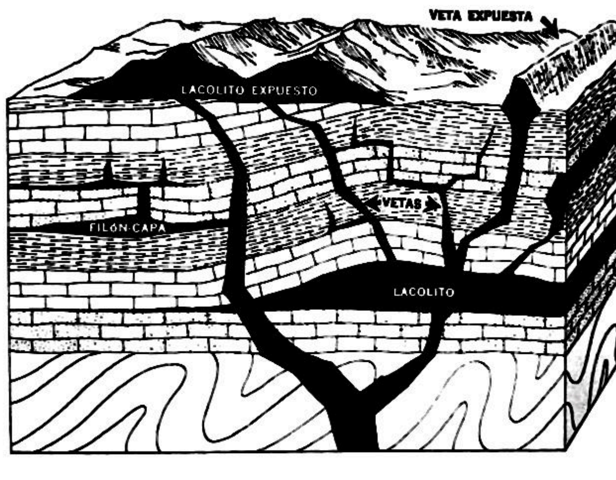
En otros casos, el magma no fue detenido en su avance ascendente por las rocas circundantes. Llegó a la superficie y fluyó, ya sea a través de una simple abertura o chimenea de volcán, o de una grieta en las rocas. Las rocas ígneas formadas por magma expelido del subsuelo se conocen como masas efusivas.

MASAS INTRUSIVAS

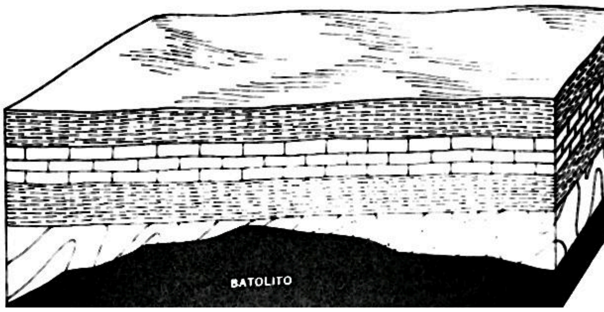
Las masas intrusivas de rocas ígneas forman una parte considerable de la corteza terrestre. A partir de su forma y de su relación con las rocas que las rodean pueden dividirse en varias clases: vetas o diques, filones-capas, lacolitos, cuellos volcánicos, batolitos y pilares.

Las vetas se formaron del magma que se abrió camino a través de las capas de roca y se solidificó en las fisuras producidas de esa manera. En algunos casos, las vetas quedan expuestas a la vista. Esto se debe a que la fisura original se extendió hasta la superficie, o a que las rocas que antes las cubrían fueron erosionadas en el trascurso de las edades. En algunas zonas, las vetas son muy numerosas y están muy próximas unas a otras. Pueden ser paralelas entre sí, o irradiar de un centro común, como los rayos de una rueda.

La masa de rocas ígneas que recibe el nombre de filón-capa está formada de manera muy similar a la veta, salvo que el magma de un filón penetra entre los intersticios de las capas de roca circundantes, es paralelo a esas capas, y desplaza los dos lechos adyacentes entre los que se insertó. Su espesor da la medida del desplazamiento de la roca circundante; el filón que resulta puede tener un grosor de muchos cientos de metros.



En el diagrama superior vemos distintos tipos de masas intrusivas. La roca fundida se desplaza por las capas de roca y forma vetas. Cuando se introduce entre dos capas produce filones; si la roca suprayacente se arquea hacia arriba, el resultado es un lacolito. El diagrama inferior muestra un batolito. Éste se produce cuando las masas intrusivas se extienden hacia abajo. Las capas de rocas que figuran en los dos diagramas están entre las identificadas a la izquierda.



Un filón puede ser horizontal o estar inclinado, junto con las capas de roca paralelas a él. En el último caso, esto puede significar que el filón se introdujo en rocas que ya estaban inclinadas o deprimidas, o en lechos planos que, junto con el filón, fueron desviados posteriormente por fuerzas de la corteza terrestre. En los Estados Unidos, los farallones conocidos como Palisades, a lo largo de la orilla occidental del río Hudson, al este de Nueva Jersey, son el borde de un filón ígneo de muchos metros de espesor que fue inclinado hacia abajo, en dirección al oeste, algún tiempo después de su formación.

A veces, durante la formación de un filón, el magma no se extiende con bastante facilidad, por lo cual la capa de roca situada sobre él se curva hacia arriba. La masa ígnea resultante, que tiene la forma de una lente vista en corte, recibe el nombre de *lacolito* (en griego, "depósito de piedra"). Tiene un piso plano, unido por un "tallo" a la fuente de magma que se encuentra debajo, de modo que, en conjunto, semeja una seta vista de lado.

La chimenea de un volcán extinguido puede llenarse con una masa cilíndrica de lava sólida, llamada *tapón* o *cuello volcánico*, que suele tener más de un kilómetro de espesor. Puede estar conectado con una gran masa subterránea de roca ígnea, llamada *de acumulación*, que quizá haya proporcionado lava al volcán en erupción durante la etapa magmática. Diques y filones emanan muchas veces del tapón. Si las rocas circundantes han sido erosionadas, el tapón sobresale como un cuello o una torre.

Las enormes masas intrusivas de roca ígnea que se ensanchan y extienden hacia abajo, hasta profundidades desconocidas, se denominan *batolitos* ("rocas profundas", en griego). Difieren de los lacolitos en que no muestran señales de un piso definido como éstos. Además, el magma del que derivan se abrió camino a través de capas de rocas y no entre ellas. Lógicamente, como en el caso de otras masas intrusivas, los batolitos quedan a la vista sólo cuando las rocas que los cubren han sido erosionadas. Uno de los más grandes del mundo forma la cadena de Sierra Nevada, en California. La erosión lo puso al descubierto en extensas superficies. Los batolitos constituyen el núcleo de muchas cadenas de montañas.

Al igual que los batolitos, las masas de acumulación son intrusivas y no tienen un piso conocido. Las dos formaciones difieren solamente en tamaño. Una masa de acumulación es un batolito que cubre una superficie de menos de 100 kilómetros cuadrados.

MASAS EFUSIVAS

Hay distintos tipos de masas efusivas. En unos casos, el magma ha sido expelido del interior de la Tierra, con fuerza explosiva, a través de una chimenea volcánica, formando una lluvia de material rocoso atomizado, expelido muy alto en el aire. Luego, los fragmentos se convirtieron en roca. En otros casos, el magma no fue arrojado desde una chimenea volcánica, sino que ha manado de ésta, derramándose como lava sobre el terreno, y se ha trasformado en roca ígnea por enfriamiento.

MINERALES DE LAS ROCAS IGNEAS

El número y variedad de los minerales que contribuyen a la formación de las rocas ígneas es más bien limitado. El más común es el cuarzo, un mineral muy duro compuesto de silicio y oxígeno. También abundan los feldespatos, que pueden ser claros u oscuros y que contienen potasio, sodio o calcio, como también aluminio, silicio y oxígeno. El piroxeno y la hornablenda, que contienen distintos metales, además de silicio y oxígeno, son más oscuros y pesados. Las micas son minerales en forma de láminas, compuestas de aluminio, otros metales, oxígeno y silicio. La magnetita, un compuesto de hierro y oxígeno, es pesada y magnética. La olivina, compuesta de hierro, magnesio, oxígeno y silicio, es un mineral verde.

DISTINTAS TEXTURAS

Las rocas ígneas difieren entre sí por su textura, vale decir, por el tamaño, forma y disposición de las partículas que las componen. El tamaño de los granos depende de la velocidad de enfriamiento del magma. Si el proceso ha sido lento, los cristales minerales han tenido mayor tiempo para formarse y por eso adquirieron un tamaño relativamente grande. El resultado es una roca de grano grueso. Pero si el magma se ha enfriado muy rápidamente, los cristales no han tenido mucho tiempo para crecer. La formación rocosa resultante es de grano fino. Por lo general, existe una relación entre la profundidad a que se ha formado la roca y el grosor de su grano; a mayor profundidad, el grano es más grueso.

En algunas rocas ígneas encontramos una cantidad relativamente grande de cristales, incluidos en una masa de otros de grano más fino, denominada matriz o pasta fundamental. Las rocas de este tipo, en las que se mezclan granos finos y gruesos, se llaman *pórfidos*.

TIPOS DE ROCAS IGNEAS

Hay muchas variedades de rocas ígneas. Algunas clasificaciones incluyen varios centenares de ellas. Sin embargo, la corteza de la Tierra está compuesta principalmente de sólo unas pocas clases. Una de las más comunes es la roca plutónica, o profunda, llamada granito, que se presenta en forma de masas enormes e irregulares. El granito está compuesto especialmente de feldespato y cuarzo. Contiene también hornablenda y mica (de la variedad negra, llamada biotita). Este tipo de roca presenta, por lo general, una textura relativamente uniforme. En las vetas se encuentra muchas veces el tipo de granito llamado pegmatita, que es de grano muy grueso. La pegmatita es origen de gran parte de la mica blanca usada en el comercio.

Al igual que el granito, la diorita es una roca de textura uniforme, pero, a diferencia de aquél, no contiene cuarzo. Está formada por feldespato y uno o más minerales oscuros, especialmente hornablenda, piroxeno y biotita. Estos materiales oscuros predominan en la roca llamada gabro, donde se encuentran también cantidades no muy grandes de feldespato. La sienita es otra formación rocosa ígnea, que no contiene cuarzo. El nombre deriva de Syene (ahora llamada Asuán), sobre el Nilo, donde hay una notable formación de esta roca.

La roca oscura y pesada llamada peridotita está compuesta íntegramente por minerales ferromagnésicos –compuestos de hierro y magnesio–. La peridotita se encuentra especialmente en masas intrusivas tales como vetas y filones. En esta formación rocosa, muy característica, se hallan depósitos considerables de níquel, cromo, platino y hierro.

Los basaltos son las rocas ígneas efusivas más comunes. Su color es negro, pardo, gris oscuro o verde oscuro. La formación basáltica de la meseta de Columbia, ubicada en la parte noroeste de los Estados Unidos, abarca una superficie de más de 500.000 kilómetros cuadrados y alcanza más de 1.200 metros de espesor. En el Decán, situado en la India occidental, y en las Islas Británicas septentrionales hay también grandes formaciones basálticas.

La riolita tiene una composición similar a la del granito. Es una roca efusiva, con un grano muy fino o felsítico, pero si tiene cristales grandes de feldespato, cuarzo o mica puede presentar una textura porfídica. La andesita es la forma efusiva de la diorita, más oscura que ésta, y puede tener también características porfídicas. Recibe su nombre por la cordillera de los Andes, situada en América del Sur, donde existe en abundancia.

Entre las más notables rocas vítreas que se encuentran en las masas efusivas están la obsi-

Philip Gendreau

U.S. Geological Survey

Arriba: Monumento Nacional Torre del Diablo en el norte de Wyoming. Lo que fue originariamente un tapón de lava sólida en la chimenea de un viejo volcán ha quedado expuesto ahora a la vista porque la erosión destruyó las rocas circundantes. Abajo: masa de lava viscosa solidificada en un volcán de Hawai.

Institución Ward de Ciencias Naturales

Institución Ward de Ciencias Naturales

Tres tipos de rocas ígneas. Arriba, izquierda: granito de biotita, compuesto de biotita (negra), feldespato (blanco) y cuarzo (gris). Arriba, derecha: riolita, roca de color claro y origen volcánico. Abajo: obsidiana, roca cristalina y muy lustrosa. Deriva del magma que se enfría rápidamente.

diana y el vidrio volcánico. La obsidiana es una roca lustrosa de color negro, gris, amarillo o pardusco. Se la encuentra en muchas partes del mundo. Una de las formaciones más grandes es el Risco de Obsidiana, en el Parque Nacional del Yellowstone, en el oeste de los Estados Unidos. El vidrio volcánico es mucho menos lustroso que la obsidiana.

ROCAS SEDIMENTARIAS

Algunas rocas sedimentarias se formaron de rocas ígneas. Las capas superficiales de estas últimas han sufrido la acción constante de las fuerzas erosivas, como cambios de temperatura, agua corriente, y también los efectos demoledores de fragmentos arrojados por el viento. A causa de esto, algunas rocas ígneas superficiales han sido fragmentadas. Estos fragmentos, barridos por los vientos, glaciares, torrentes y corrientes marítimas costeras se depositaron en lagos o en partes poco profundas del mar. Presionados por el peso de acumulaciones posteriores, se trasformaron gradualmente en rocas sedimentarias. Éstas también se formaron con restos de plantas y animales y por evaporación del agua de mar.

Las capas sedimentarias de la corteza terrestre constituyen la mayor parte de su superficie, y el grosor de cada una de ellas oscila entre unos pocos centímetros y varios metros. En algunos lugares se encuentran sólo unas pocas capas. En otros hay grandes acumulaciones de lechos, que llegan a tener un espesor total de varios kilómetros.

Hablando en términos generales, puede decirse que cuanto más antiguo es el lecho, tanto más completa es la cementación de las rocas sedimentarias. Los fragmentos de rocas nuevas tienen tan poca cohesión que algunas veces pueden ser extraídos cavando con una pala. Pero las rocas más viejas son generalmente tan sólidas que resulta difícil extraerlas, salvo que se las barrene, se las vuele con dinamita y se las rompa en fragmentos todavía más pequeños con un pico o un martillo de clivaje.

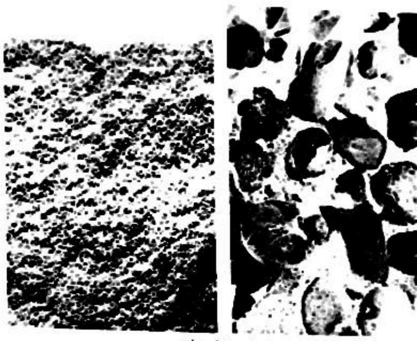
ROCAS DETRÍTICAS O CLÁSTICAS

Las rocas sedimentarias, formadas íntegramente por partículas de otras, son conocidas como rocas detríticas o clásticas (del griego klastos, que significa "roto"). Los fragmentos de los que derivan las rocas clásticas se clasifican, partiendo de su tamaño, en cuatro grupos: grava, arena, limo y cieno.

La grava es el más grueso de los sedimentos. Consiste en fragmentos que tienen como mínimo dos milímetros de diámetro. El tamaño de los fragmentos de grava va desde pequeños cantos rodados hasta grandes pedrejones. Entre estos extremos se encuentran los guijarros. Los sedimentos de arena consisten en granos redondeados que van desde 2 milímetros hasta 1/16 de milímetro de diámetro. Las partículas de menos de 1/16 de milímetro, y por lo menos 1/256 de milímetro de diámetro se llaman limo. Los sedimentos más finos –cienos y arcillas– están formados por fragmentos de diámetro inferior a 1/256 de milímetro.

W. A. Schwartz

Institución Ward de Ciencias Naturales



Ambas fotos: Institución Ward de Ciencias Nat

Cuatro tipos de roca sedimentaria. Arriba, izquierda: evaporita (nitrato de sodio); abajo, coral. Arriba, derecha: piedra arenisca; extrema derecha: conglomerado (pudinga).

Cuando algunas partículas más o menos redondeadas de grava –piedrecitas, guijarros y cantos rodados– se consolidan, forman la roca llamada *conglomerado*. En la variedad conocida como brecha, los fragmentos de grava son angulosos y no redondeados.

Los granos de arena cementados originan la formación porosa conocida como piedra arenisca. Los poros de esta roca pueden constituir hasta el 30 por ciento del volumen total, y los líquidos pasan libremente a través de los mismos. Por esta razón, las piedras areniscas muchas veces constituyen depósitos de petróleo o de agua subterránea. Las partículas de limo se convierten en piedras de limo, y las de cieno, en piedras de cieno y pizarra.

ROCAS DERIVADAS DE PLANTAS Y ANIMALES

Las rocas más comunes derivadas de plantas son las distintas variedades de carbón. Todo el carbón comenzó siendo material verde, o leñoso, restos de plantas que prosperaron hace muchos años en zonas pantanosas. Estos restos quedaron sepultados bajo sedimentos posteriores. Los distintos estratos se hundieron debajo de antiguos mares, pero surgieron periódicamente. Algunos fueron plegados y fragmentados por los movimientos de la corteza. De resultas de la tremenda presión, el calor y la exclusión de aire, con el correr del tiempo los materiales orgánicos se trasformaron en carbón. La turba, el lignito, la hulla y la antracita representan estadios sucesivos de esta trasformación.

Los restos animales contribuyeron a la formación de distintos tipos comunes de rocas sedimentarias. Islas y arrecifes coralíferos están constituidos por los esqueletos de innumerables pólipos de coral y otros organismos. Una generación de coral tras otra ha contribuido a formar esos enormes depósitos.

La roca llamada piedra caliza está compuesta de calcita, o carbonato de calcio, un mineral formado por calcio, carbono y oxígeno. La calcita se encuentra comúnmente en los caparazones de animales como caracoles acuáticos, almejas, braquiópodos y otros mariscos. Estos caparazones pueden contribuir a la formación de piedra caliza. Sin embargo, gran parte de esta formación rocosa está formada por carbonato de calcio, precipitado directamente en el agua. Hay piedras calizas marinas y fluviales. Cuando las aguas se retiran, los depósitos de piedra caliza se convierten en parte de las rocas terrestres.

EVAPORITAS

Algunos depósitos sedimentarios han sido formados por sales precipitadas, o separadas por solución, en agua que se evapora. Estos sedimentos se llaman evaporitas. Pueden desarrollarse a partir del agua del mar, o de lagos y torrentes. Muchas de las sales consideradas se encuentran tanto en las aguas oceánicas como en las terrestres. Sus concentraciones pueden diferir, y los depósitos minerales que forman varían en consecuencia. Las evaporitas señalan la existencia de condiciones climáticas cálidas y secas. Un brazo de mar puede ser aislado y trasformado en un lago que llega a tener una gran concentración de sal a medida que el agua se evapora. En las regiones desérticas, las corrientes y charcos temporarios pueden ser soluciones saturadas de sal que dejan depósitos si el sol los evapora.

Cuando comienza la evaporación, los compuestos menos solubles precipitan antes. En depósitos derivados de una masa de agua de mar, el carbonato de calcio y el óxido de hierro forman la capa primera, o inferior. Luego viene el yeso, un compuesto de calcio, que contiene azufre y oxígeno (sulfato). La sal común, o sal de roca, precipita inmediatamente después. La siguen otros compuestos, tales como los sulfatos de magnesio y sodio. Cuando estos cuatro grupos de minerales disueltos se encuentran juntos, forman capas sucesivas de evaporitas desde el fondo hacia arriba. Si las condiciones de sequedad se prolongan, los depósitos subsisten. De otra manera, la lluvia, o las aguas de reflujo, los disolverán nuevamente, o los cubrirán con capas de cieno o de arena.

El nitrato de sodio, un compuesto de sodio que contiene oxígeno y nitrógeno, forma grandes depósitos de evaporita en los desiertos del norte de Chile. Este nitrato fue producido, originariamente, por ciertas bacterias de las plantas. Luego el agua inundó la región y disolvió el nitrato del suelo. Tiempo después, cuando las aguas se secaron, quedó este compuesto.

FÓSILES

A medida que se acumulan los sedimentos, pueden contener restos de organismos muertos. Éstos se convierten en piedra o dejan sus improntas en las rocas formadas por sedimentos y de esta manera se trasforman en fósiles. Los científicos los estudian para descubrir la evolución de la vida a través de las edades. Todas las rocas sedimentarias fosilíferas conocidas, dispuestas capa por capa, forman lo que se conoce con el nombre de *columna estratigráfica*. Ésta es la medida histórica más usada por los geólogos para fechar los acontecimientos geológicos.

Ejemplos de roca metamórfica. De arriba hacia abajo: pizarra negra; gneis de biotita; cuarcita; mármol grafítico.

ROCAS METAMÓRFICAS

La tercera gran división de las rocas, las rocas metamórficas, se compone, como ya lo hemos señalado, de rocas ígneas o sedimentarias trasformadas dentro de la corteza terrestre en otras de un tipo muy diferente. Pero, por lo general, han preservado algunos rasgos de su estructura original.

FORMACIÓN

Uno de los factores más importantes en el metamorfismo –formación de roca metamórfica– es la presión. Puede ser provocada por los lechos sedimentarios suprayacentes, originada por el magma que se abre camino dentro de las capas de roca circundantes, o ser el resultado de fuerzas tectónicas que deforman la corteza terrestre. Cuando actúa una presión se genera enorme calor. Éste acelera las reacciones químicas que se producen e intensifica sus efectos. La presencia de agua es otro factor de metamorfismo. El agua suele ser tan escasa que forma sólo una película alrededor de las partículas. Empero, proporciona un medio donde las sustancias de la roca pueden entrar en solución y luego condensarse sobre la superficie de nuevos cristales en crecimiento.

Bajo la acción de presión y calor extremos, las partículas de la roca original se ven forzadas a adoptar nuevas disposiciones. En algunos casos, los componentes de la roca combinan con otros que se encuentran en la vecindad inmediata y forman nuevos minerales, muchos de los cuales se desarrollan en forma cristalina casi perfecta, como el granate o almadina.

ESTRATIFICACIÓN

Las rocas metamórficas suelen mostrar capas semejantes a las de las sedimentarias. La estratificación metamórfica, o foliación, se debe a cambios mecánicos o químicos producidos en la roca original. Los granos, cristales y fósiles son desviados o quebrados y extendidos en series lineales. Pueden formarse filas paralelas de minerales en forma de placas, que no están relacionados para nada con la estratificación original. Por obra de la presión, la lutita se trasforma en pizarra. Ésta, a su vez, puede convertirse luego en una roca llamada filita y, bajo una presión continua, en una roca cristalina foliada, conocida como esquisto. Otros tipos de roca, como los granitos y pizarras arenosas, se trasforman en una roca cristalina de foliación más gruesa, el gneis.

La pizarra es una roca oscura, con grano fino invisible. Se parte fácilmente en lajas finas y suaves. La filita representa un estado de transición entre la pizarra y el esquisto. Su grano es muy fino, compuesto especialmente de mica, pero en distintos lugares aparecen unos pocos cristales más grandes. La foliación es más tosca y ondulada que la de la pizarra. El esquisto es visiblemente cristalino, con foliación ondulada. Consiste en minerales escamosos o tabulares, como mica, clorita, hornablenda o talco, y también algunos cristales de cuarzo y granate.

El gneis tiene aspecto irregular y veteado a consecuencia de las capas alternadas de distintos minerales y se origina a partir de diferentes clases de rocas. Cada tipo tiene casi la misma composición que la roca madre. El granítico, por ejemplo, deriva del granito.

Otros tipos de rocas metamórficas no muestran ninguna foliación. Las piedras areniscas puras se trasforman en una masa más compacta llamada cuarcita, donde los espacios porosos han sido comprimidos hasta formar una roca muy dura y durable. El calor y la presión trasforman la piedra caliza en mármol, donde los granos de carbonato se vuelven visiblemente cristalinos. El mármol puro es blanco. Sin embargo, la mayor parte de las piedras calizas contiene impurezas que muchas veces reaccionan durante el metamorfismo y producen nuevos minerales. Éstos suelen dar al mármol resultante colores llamativos o jaspeados, haciéndolo más atractivo como material para construcción. Sometido a presión, el mármol es muy plástico.

METAMORFISMO INVERTIDO

También suele producirse el metamorfismo invertido. Cuando las fuerzas trasformadoras dejan de actuar, las rocas metamórficas tienden a recuperar su condición original inalterada; pueden volver a fundirse en magmas. Todo el ciclo comienza entonces nuevamente, porque los procesos formadores de rocas antes descritos continúan actuando. Rocas nuevas reemplazan a las destruidas por los efectos del agua y del aire, prosiguiendo así un ciclo geológico continuo.

DERIVA CONTINENTAL

La noción de que el suelo puede trasladarse es muy antigua y ha hallado su expresión en el folklore y la mitología; pero sólo en el siglo XX se propuso como teoría científica seria. En resumen: deriva continental significa que continentes enteros se han desplazado, o trasladado, miles de kilómetros en períodos de millones y, quizá, miles de millones de años. Esto probablemente ocurrió no sólo una vez, sino varias, durante el pasado geológico. Las masas terrestres cambiaron su posición unas respecto de otras y de los polos geográficos, a través de los tiempos, y pueden continuar haciéndolo. Se han unido y separado muchísimas veces. La velocidad de movimiento quizá sea tan sólo de unos pocos centímetros por año.

Algunos teóricos de la deriva sostienen que los continentes flotan como barcos en un mar de rocas plásticas. Otros piensan que las masas continentales están estancadas en una corteza que se desliza de varias maneras, arrastrando las masas en forma pasiva, como troncos atascados entre hielos flotantes. Los continentes pueden haber seguido cualquiera de las muchas trayectorias posibles. Pocas autoridades coinciden respecto a cuáles; tampoco hay acuerdo respecto de las velocidades de movimiento, las épocas en que los continentes se reunieron o separaron, y cómo ocurrieron estos sucesos.

Cuando se propuso por primera vez la idea de deriva continental, se sabía mucho menos que ahora sobre la Tierra; en consecuencia, muchas de las ideas iniciales ya han quedado desacreditadas. Por añadidura, casi desde el comienzo, el concepto de continentes erráticos se topó con la rígida oposición de algunos geólogos conservadores. La estructura compleja del conocimiento geológico moderno se asienta en la doctrina conocida como permanencia de los continentes y cuencas oceánicas, según la cual la corteza terrestre es bastante estable, y la ubicación actual de continentes y océanos fue determinada muchísimas edades atrás. Sin embargo, la escuela de la permanencia acepta la existencia de acontecimientos locales, o comparativamente menos radicales, como el ascenso y hundimiento del suelo o del mar, orogénesis, terremotos, edades glaciales, volcanes y demás.

Pero se ha observado, como veremos más adelante, que la doctrina de la permanencia ignoraba –o desconocía– ciertos problemas en la distribución de rocas, fósiles y seres vivientes sobre la faz de la Tierra; o bien los explicaba sin recurrir a la deriva continental. La escuela de la permanencia aún se considera con respaldo, y ha ofrecido serias objeciones a la deriva continental.

HISTORIA DE LAS TEORÍAS SOBRE DERIVA

Poco después del trazado cartográfico del océano Atlántico –hace unos cuatrocientos años– los geógrafos advirtieron que las costas occidentales de Europa y África parecían coincidir con las opuestas del este de ambas Américas. Acercando estos continentes, por así decir, se los podía hacer coincidir a lo largo de estas costas. Pero sólo en 1858 se publicó un mapa que demostraba cuán bien coincidían. Se sugirió que estos continentes habían estado antes unidos. Esta vez la idea prendió. Hacia 1900, varios geólogos propusieron, también, que ciertas tierras del hemisferio austral –grandes zonas de África, Madagascar, India, Australia, Antártida y América del Sur– estuvieron una vez unidas en un enorme supercontinente llamado *Gonduana*, por la región de Gonduana, en India central. Sin embargo, la gran mayoría de los geólogos conservadores rehusó tomar estas ideas muy en serio.

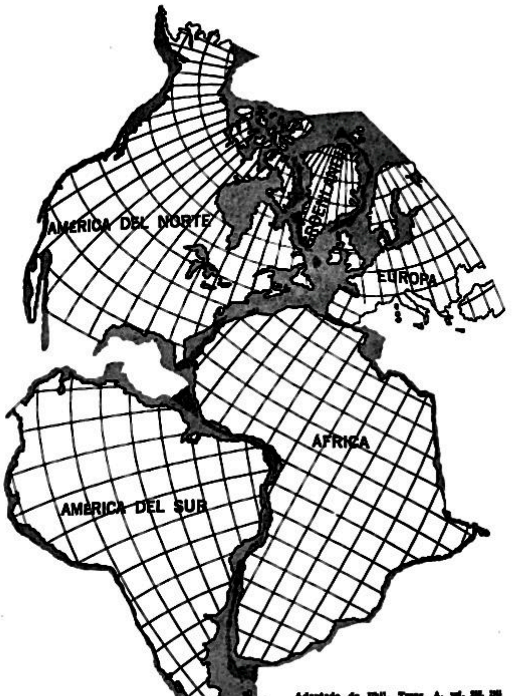
Pero entre 1912 y 1920, el científico alemán Alfred Wegener ofreció una teoría importante sobre la deriva continental; propuso que, en un momento de la historia de la Tierra, entre 200 y 300 millones de años atrás –en el Paleozoico Superior– todos los continentes actuales estaban unidos en una sola inmensa masa continental llamada *Pangea* (en griego, "toda la Tierra"). En ese momento, la Pangea comenzó a desmembrarse lentamente, dividiéndose en los continentes que la componían; éstos, con el tiempo, se desplazaron hasta sus posiciones actuales. Según Wegener, la fuerza propulsora de este proceso actuó en direcciones divergentes de los polos geográficos. Se dijo que fue el resultado de fuerzas gravitatorias e inerciales relacionadas con la rotación de la Tierra en torno de su eje. Después se demostró que esa fuerza es demasiado débil para ejercer efectos apreciables sobre la corteza terrestre. Wegener también concibió a los continentes derivando fácilmente sobre la roca subyacente, la que cedía lentamente a la presión. Esta idea también goza de poco crédito hoy en día, incluso entre los derivistas. Wegener sostuvo que Europa y Groenlandia se han estado separando en tiempos históricos, y basó su argumento en mediciones sucesivas de latitud y longitud. Después se halló que estas determinaciones geográficas eran muy imprecisas para tener algún valor en la medición de la deriva.

Pero muchas otras ideas de Wegener son sólidas, y han continuado aportando una base firme para ulteriores teorías sobre deriva. En verdad, varias se habían formulado incluso antes de que escribiera Wegener. Se señaló que muchos lechos pétreos, estructuras rocosas (tales como pliegues y fallas, o grietas), fósiles e incluso organismos vivos eran tan parecidos en continentes ampliamente separados, que no podían haberse originado en forma independiente. Según los teóricos de la deriva, al principio deben haberse desarrollado juntos en una sola masa terrestre que luego se desmembró. Tampoco las plantas y animales pudieron haber emigrado a través de barreras oceánicas. Asimismo, los teóricos de la deriva continental sostienen que las antiguas zonas climáticas –tal como se han determinado a partir de las características de rocas y fósiles– no corresponden con las modernas, lo que sugiere que, en otros tiempos, los continentes estuvieron en latitudes diferentes, inferencia que niegan los geólogos antiderivistas.

de Alfred L. Wegener, El origen de continentes y océanos, Dover Publications, Inc N.Y. Reimpreso con autorización del editor

La deriva continental, tal como fue reconstruida por Alfred Wegener. En el planisferio del extremo izquierdo (página anterior) vemos todos los continentes actuales unidos en una sola masa continental gigantesca, llamada Pangea por Wegener, en una época de la historia de la Tierra conocida como Paleozoico Tardío, alrededor de 250 a 300 millones de años atrás. Hacia principios del Cenozoico, hace unos 60.000.000 de años, tal como se muestra en el mapa del medio, Pangea se había desmembrado tanto que los continentes estaban parcialmente separados entre sí. Durante la época del Pleistoceno (era Cenozoica muy tardía, cerca de 1.000.000 de años atrás) los continentes ocuparon posiciones muy similares a las que presentan hoy. Los geólogos modernos sostienen que, en cambio, hubo originalmente dos enormes supercontinentes. En estos mapas, las zonas blancas son el océano profundo; las grisáceas claras, masas continentales, y las grisáceas oscuras, mares poco profundos.

Este mapa muestra cuán bien coinciden los bordes occidentales de Europa y África, con los orientales de América del Norte y del Sur. Esta reconstrucción fue hecha por sir Edward Bullard, geofísico inglés, y otros, tomando como base la profundidad de 1.000 metros de las plataformas continentales, que son los bordes verdaderos de los continentes, bajo los océanos. Este método ha proporcionado un ajuste aún mejor que el basado en las mismas líneas costeras de estos continentes. Si bien el mapa indica que los continentes formaron alguna vez una sola masa, muchos geólogos creen que partes de África, América del Sur y otras tierras australes estuvieron unidas en un gran continente separado –Gonduana– y partes de América del Norte, Europa y Asia, en otro –Laurasia–. Las áreas sombreadas que circundan las tierras son las plataformas continentales; las oscuras, los lugares donde se superponen plataformas opuestas.



Adaptado de Phil. Trans. A, vol. [illegible]
Consejo de la Royal Society

La batalla causada por las ideas de Wegener amenguó hacia 1930, con la victoria para la escuela de la permanencia, pero en la década del 30 varios geólogos todavía se aferraban a los conceptos de Wegener; estos científicos provenían principalmente de las tierras del hemisferio austral, donde la evidencia de la deriva es más convincente. La autoridad más prestigiosa era Alex L. du Toit, de Sudáfrica, para quien no hubo uno, sino dos supercontinentes primitivos en este planeta: Gonduana en el hemisferio sur, y *Laurasia*, llamado así por las montañas Laurentinas del Canadá, más Asia, en el hemisferio norte. Ambos supercontinentes estaban separados por un amplio mar llamado *Tethys*. Laurasia consistía en partes de América del Norte, Groenlandia, Europa y Asia. Ambos supercontinentes se desmembraron con el tiempo. Todavía es arduo el problema de cómo discernir Laurasia y Gonduana. Algunas autoridades incluirían a Nueva Zelanda en Gonduana, y también trasferirían las partes boreales de América del Sur y Africa, de Gonduana a Laurasia.

Desde la década de 1940 se han hecho progresos enormes en numerosos campos del estudio de la Tierra. Ahora parece que nuestro mundo es sumamente inestable y dinámico, hasta un punto nunca percibido antes por la escuela de la permanencia. En consecuencia, se ha revivido la teoría de la deriva –particularmente desde la década de 1950– y, en la actualidad, está atrayendo a una cantidad creciente de científicos. El mayor obstáculo para la aceptación de la deriva continental ha sido la aparente carencia de un mecanismo suficientemente verosímil que la produzca. En apariencia, este mecanismo es hoy comprensible y se lo explica según una teoría conocida ahora como *tectónica de placas*. Empero, el problema no está totalmente resuelto aún.

MECANISMOS DE DERIVA CONTINENTAL

La Tierra es casi tan rígida como una bola de acero en su reacción a breves choques. ¿Cómo podrían entonces trasladarse partes enteras de su superficie miles de kilómetros? La razón es que la Tierra se mueve lentamente y varía en su estructura y composición. Tiene una capa externa delgada –la corteza– que contiene granito en los continentes y lechos más delgados de basalto bajo los océanos. Por debajo de la corteza, y extendiéndose a considerable profundidad, se halla el manto, compuesto por roca diferente de la que forma la corteza. En torno del centro de la Tierra se halla el núcleo metálico.

Los experimentos han demostrado que los materiales sólidos –inclusive rocas– pueden deformarse o realmente fluir bajo elevadas presiones y temperaturas. Tales condiciones prevalecerían incluso a porfundidades moderadas de la Tierra, tanto en la corteza como por debajo de ella. También se produciría fusión o fractura de la roca. Estos fenómenos ayudarían a explicar muchas clases de movimientos y actividades de la Tierra, como plegamiento de lechos rocosos, orogénesis, ascenso y hundimiento del terreno, vulcanismo, terremotos, y, por supuesto, deriva continental.

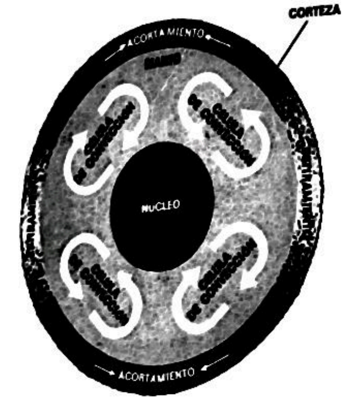
Los primeros teóricos pensaron que los movimientos planetarios de la Tierra podrían ser origen de importantes cambios corticales en la deriva continental. Los que vinieron luego propusieron ideas tales como las del deslizamiento de los continentes hacia depresiones de la corteza, o la expulsión de masas de tierra como barajas fuera de un mazo. Pocas de esas ideas se toman en serio hoy en día.

Se necesitaba un conjunto de fuerzas mucho más poderoso para hacer que se trasladaran los continentes, y los geólogos lo buscaron en el interior de la Tierra, donde se considera que hay un calor muy intenso, causado por la profusión de elementos radiactivos que emiten radiación nuclear y calientan lentamente las rocas circundantes. Asimismo, el terrible peso de la roca suprayacente comprime el material rocoso inferior, calentándolo más y permitiéndole, inclusive, fluir. Tal como sabemos, el calor es una forma muy poderosa de energía que debe ejercer algún efecto sobre las rocas.

Células de convección. Algunos geólogos creen que la roca calentada reacciona en forma muy parecida a la de una masa de agua o gas calentados. La mayoría de nosotros ha visto agua puesta a hervir en un recipiente; a medida que se caldea, la más caliente –por ser menos densa que la circundante más fría— asciende a la superficie, pierde algo de calor, se hace más densa, y vuelve a hundirse hacia el fondo del recipiente. Este ascenso y descenso repetidos de las corrientes de agua provocado por el calor se llama *convección*. El recorrido de una corriente, o conjunto de ellas, es aproximadamente circular y suele denominarse célula de convección. Según la naturaleza, volumen, forma y temperatura del material afectado, puede formarse en su seno cualquier número de células de convección. La convección es un medio para trasferir energía calórica.

Un grupo considerable de geólogos piensa que la roca del manto terrestre puede desarrollar células de convección similares por calentamiento, el que no es uniforme por la desigual distribución de elementos radiactivos en el interior. Tales corrientes de convección se moverían muy lentamente, porque la roca es sólida. Sin embargo, bajo calor y presión extremos, ésta puede comportarse como un semisólido o un líquido sumamente espeso y viscoso. Los geólogos han hallado justificación para esta idea en la conducta de las cúpulas o tapones de sal, que son enormes masas de sal gema que, en el trascurso de millones de años, se han abierto camino literalmente, a través de miles de metros, desde las profundidades en la Tierra y al hacerlo han combado y triturado las rocas que las rodeaban. Las autoridades atribuyen este fenómeno a que la sal es menos densa que las rocas circundantes, a través de las cuales ha ascendido tal como un globo con helio sube en el aire más denso, haciéndolo a un lado.

Fractura de la corteza. La teoría de la convección rocosa se emplea para explicar la deriva continental del modo siguiente: a medida que una corriente de convección asciende ha-



Células de convección en el manto rocoso del interior de la Tierra. Pueden ser la causa principal de la deriva continental. A medida que la roca calentada asciende y se hunde lentamente en circuitos aproximadamente circulares, o células, los movimientos combinados ocasionan el estiramiento, o acortamiento, de la corteza suprayacente. Allí donde las corrientes rocosas se aproximan entre sí, a partir de células adyacentes, la corteza es acortada y plegada, para formar montañas y arcos insulares. Donde las corrientes se alejan unas de otras, la corteza se estira y se fractura: se forman grietas y cordones, y se agrega nueva roca a la corteza.

cia la parte superior del manto, choca contra la base de la corteza fría y quebradiza, desviándose a lo largo de la misma. A causa de la fricción, la corteza se ve arrastrada, junto con la ahora horizontal corriente del manto, y desplazada. Donde existen varias células gigantes de convección que perturban así la corteza, ésta se distorsiona y fractura en grandes pedazos, o placas, que se separan.

Donde convergen corrientes provenientes de un par de células de convección adyacentes, cuando vuelven a descender al interior de la Tierra arrastran consigo a la corteza, hasta cierta profundidad. Si esto sucede bajo un continente, el suelo es arrastrado y combado simultáneamente, formando con el tiempo un plegamiento montañoso, tal como los Andes. Si las corrientes rocosas descendentes convergen bajo el piso oceánico, se produce un suceso similar, pero con resultados algo diferentes; en el fondo del océano se forma una depresión profunda, con arcos de islas montañosas escarpadas en su borde. Estas fosas submarinas con sus arcos de islas, existen en la actualidad, especialmente frente a las costas del Pacífico de América del Norte y Asia.

Expansión del fondo del mar. Donde las corrientes rocosas ascienden y divergen desde un par de células de convección adyacentes, la corteza, en cambio, se estira y fractura. Si esto ocurre debajo de un continente, éste puede fragmentarse. Bajo el lecho oceánico, en cambio, ocasiona una fractura –o grieta– en su corteza delgada. En todos los fondos oceánicos existe un sistema mundial de tales grietas sumergidas, con sus cadenas montañosas correspondientes, lo cual se toma como evidencia de estiramiento y fractura cortical en esas zonas; una expansión del fondo del mar. Se considera que las regiones de montañas y de arcos insulares son sitios de contracción y absorción corticales, particularmente en los bordes del océano Pacífico. Las zonas de grietas oceánicas son lugares donde la corteza se está estirando y donde, de hecho, se forma nueva corteza. A medida que el antiguo material cortical se aparta de cada lado de la grieta, surge del manto roca que reemplaza a la más vieja y se convierte en parte de la corteza.

Probablemente, todo el proceso no es continuo; existe un límite para la cantidad de material cortical que puede ser arastrado hacia abajo y absorbido por el manto, que es más denso que la corteza. El resultado final de este crecimiento cortical, estiramiento y acortamiento es que los continentes se trasladan de un punto a otro, desmembrándose y reuniéndose a través de las edades. De este modo, sus bordes se plegaron para formar cordilleras, mientras que el interior permanecía, con frecuencia, relativamente sereno.

Algunas autoridades creen que este mecanismo –el concepto de tectónica de placas– no es suficiente para causar movimientos continentales importantes; otra sugerencia es que la Tierra ha estado –y está continuamente– en expansión. Por ello, los continentes siguen separándose unos de otros de manera parecida a las marcas pintadas en la superficie de un globo cuando se lo infla. Pero esta idea no se ha comprobado con evidencias de expansión, ni explica cómo unos continentes se desplazaron hacia otros y se fusionaron con ellos.

PRUEBAS DE LA DERIVA CONTINENTAL

Al examinar las pruebas en pro y en contra de la teoría de la deriva, los científicos hacen uso de información obtenida en casi todos los campos de la geología y ciencias relacionadas. Tomaremos uno por vez, y evaluaremos qué datos presenta en este caso.

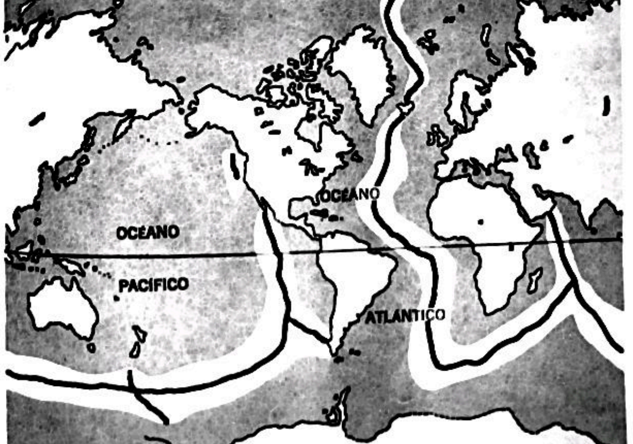
Geografía. Ya hemos hablado de la concordancia de las costas atlánticas de Europa y África, por un lado, y las de América del Norte y Sur por otro. Empero, los científicos han observado que las plataformas submarinas de estos continentes coinciden aún mejor. Éstas representan sus verdaderos límites. Se seleccionó la profundidad de 1.000 metros.

También se ha logrado la concordancia de las otras líneas costeras en el hemisferio austral, para la reconstrucción de Gonduana. Pero, como hemos dicho, no fue tan fácil concordar a ésta con Laurasia. Es muy probable que hayan sido masas terrestres diferentes durante gran parte del tiempo geológico.

Estratigrafía. La estratigrafía trata de la naturaleza y distribución de las capas o lechos rocosos, y también se interesa en su contenido de fósiles (si lo hay) y en su datación. Ciertas capas de continentes ahora muy separados son de naturaleza tan similar, que muchos estratígrafos piensan que deben de haberse formado al mismo tiempo en una única zona general, cuando esos continentes estaban unidos; el ejemplo más notable es la llamada serie de lechos Gonduana, en América del Sur, África del Sur, Antártida, India y Australia. Consisten en espesores variables de carbón y depósitos glaciales, con restos de plantas terrestres antiguas parecidas a helechos,

Sistema mundial de grietas y cordones oceánicos. Ocupa, aproximadamente, las regiones centrales de los fondos del mar y consiste típicamente en zonas lineales deprimidas con cordones montañosos que corren por su medio. Las grietas también pasan por zonas continentales, notablemente en África oriental (no se muestra), Islandia y California. Estas grietas y cordones pueden ser lugares de estiramiento y fractura de la corteza, la que se aleja a cada lado de un cordón, mientras surge nueva roca para reemplazarla. Como compensación, acortamiento y absorción corticales ocurren en otras partes, particularmente a lo largo de las zonas de montañas y arcos insulares, tales como Asia oriental, Islas Aleutianas frente a Alaska, oeste de América del Sur, y otros lugares. Enormes fracturas corticales (no se muestran) intersecan los cordones.

Adaptado de Enciclopedia Harper de la Ciencia, editada por James R. Newman (Harper & Row, 1963, p. 780)



conocidas como *Glossopteris* y *Gangamopteris*. Su edad es de unos 200 a 300 millones de años. Los geólogos creen que estas capas fueron depositadas antiguamente en el continente único de Gonduana, opinión reforzada por el hecho de que las plantas terrestres no pudieron haber cruzado por sí mismas anchas barreras oceánicas para diseminarse por continentes distintos.

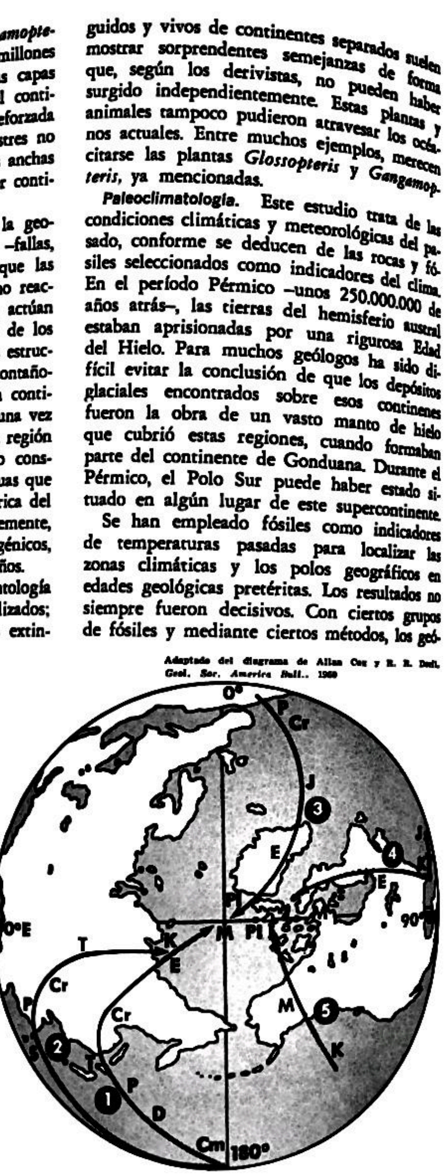
Geografía estructural. Este ramo de la geología trata de las estructuras naturales –fallas, pliegues, cuencas, cúpulas y demás– que las rocas de la Tierra pueden formar como reacción contra las poderosas fuerzas que actúan sobre ellas. Lo mismo que en el caso de los lechos rocosos, pueden rastrearse ciertas estructuras de roca –como plegamientos montañosos– por miles de kilómetros desde un continente al siguiente, lo cual sugiere que una vez estuvieron unidas y se formaron en una región ininterrumpida. Un ejemplo notable lo constituyen las estructuras montañosas antiguas que pueden rastrearse entre el este de América del Norte y el oeste de Europa. Aparentemente, fueron creadas en varios episodios orogénicos, hace varios centenares de millones de años.

Paleontología y biología. La paleontología trata de los organismos antiguos fosilizados; la biología, de los vivientes. Los seres extinguidos y vivos de continentes separados suelen mostrar sorprendentes semejanzas de forma que, según los derivistas, no pueden haber surgido independientemente. Estas plantas y animales tampoco pudieron atravesar los océanos actuales. Entre muchos ejemplos, merecen citarse las plantas *Glossopteris* y *Gangamopteris*, ya mencionadas.

Paleoclimatología. Este estudio trata de las condiciones climáticas y meteorológicas del pasado, conforme se deducen de las rocas y fósiles seleccionados como indicadores del clima. En el período Pérmico –unos 250.000.000 de años atrás–, las tierras del hemisferio austral estaban aprisionadas por una rigurosa Edad del Hielo. Para muchos geólogos ha sido difícil evitar la conclusión de que los depósitos glaciales encontrados sobre esos continenes fueron la obra de un vasto manto de hielo que cubrió estas regiones, cuando formaban parte del continente de Gonduana. Durante el Pérmico, el Polo Sur puede haber estado situado en algún lugar de este supercontinente.

Se han empleado fósiles como indicadores de temperaturas pasadas para localizar las zonas climáticas y los polos geográficos en edades geológicas pretéritas. Los resultados no siempre fueron decisivos. Con ciertos grupos de fósiles y mediante ciertos métodos, los geólogos derivistas han hallado zonas climáticas y posiciones de los polos notablemente diferentes de las actuales, y por tanto arribaron a la conclusión de que los continentes una vez estuvieron en otras latitudes. También señalan la presencia de depósitos de edades de hielo en regiones ahora soleadas de África del Sur y América del Sur, y el hallazgo de plantas subtropicales y restos animales en la extensión hoy congelada de las regiones polares. Pero otros investigadores, trabajando también con fósiles y rocas, no han confirmado estas ideas; en cambio, alegaron haber demostrado que las zonas climáticas y los polos estaban situados en prácticamente el mismo lugar que tienen ahora, y atribuyen las variaciones de los climas antiguos a factores convencionales atmosféricos, fisiográficos y solares.

Adaptado del diagrama de Allan Cox y R. R. Doell, *Geol. Soc. America Bull.*, 1960



Movimientos del polo norte magnético en un lapso de casi 600.000.000 de años, según se ha determinado a partir del paleomagnetismo de rocas antiguas (ver abajo la lista de números y letras para clave del mapa). Los nombres de los períodos y épocas corresponden a tiempos definidos en la historia de la Tierra: el primero, el cámbrico, de 500 a 600 millones de años atrás; el último que se da aquí: la época del plioceno, alrededor de 1.000.000 a 2.000.000 de años atrás. Las trayectorias variables de los polos representadas para diferentes áreas continentales indican que los movimientos de éstas ocasionaron gran parte del nomadismo polar.

(1) Europa (2) América del Norte (3) Australia (4) India (5) Japón.

Cm, Cámbrico; S, Silúrico; D, Devónico; Cr, Carbonífero; P, Pérmico; T, Triásico; J, Jurásico; K, Cretácico; E, Eoceno; M, Mioceno; Pl, Plioceno.

Geocronología. La geocronología trata de la datación de rocas y fósiles según varios métodos. Una confirmación aparentemente importante de la deriva continental es la datación de ciertos grupos de rocas similares en el norte del Brasil y en África occidental, mediante su contenido en minerales radiactivos. Dos series de roca –ambas presentes en las regiones mencionadas– han sido datadas en 2.000 y 550 millones de años, respectivamente, lo que indicaría que se formaron cuando América del Sur y África estaban unidas.

Geofísica. Es un tópico sumamente amplio, que abarca la física de la Tierra: calor, electricidad, magnetismo, procesos nucleares, gravedad, movimiento, terremotos y demás. Se han deducido pruebas de la deriva, ante todo del estudio del magnetismo terrestre, o geomagnetismo, del calor y los movimientos de nuestro planeta, y de la geocronología, en relación con la geofísica.

Científicos como Wegener intentaron medir el supuesto movimiento de deriva continental, comparando determinaciones de latitud y longitud antiguas y recientes. Esto falló, porque las mediciones nunca fueron lo bastante exactas. Empero, recientemente se idearon muchos instrumentos nuevos con la precisión requerida; algunos se usan para captar y registrar los movimientos reales de la Tierra, por leves que sean. Se han aplicado, usualmente, al estudio de terremotos, actividad volcánica y movimientos a lo largo de fallas. Otros instrumentos pueden indicar posiciones en la superficie de la Tierra, con precisión de hasta unos pocos metros, algo nunca antes logrado. En esto también colaboraron los satélites artificiales terrestres. Cualesquiera cambios ulteriores en la posición de estos puntos pueden ser pruebas de deriva continental.

En función del problema de la deriva, los estudios geotérmicos se concentran en los flujos de calor provenientes del interior de la Tierra, lo que indicaría el movimiento de corrientes de roca. Un flujo de calor elevado significa corrientes ascendentes; otro débil, corrientes descendentes. En las grietas mesoceánicas se han obtenido valores elevados de calor, tal como lo anticipaba la teoría, puesto que son regiones de corrientes ascendentes y corteza en estiramiento. Pero también se han medido a veces valores elevados de calor en las zonas de montañas y depresiones insulares, si bien la teoría pronosticaba valores bajos de calor por las corrientes rocosas descendentes. Esto no asesta un golpe serio a la teoría de la deriva en su totalidad, pero arroja cierta duda sobre la hipótesis de la convección de rocas como causa de la deriva. Quizá tenga, pues, que revisarse la teoría corriente.

La Tierra posee un campo magnético muy débil, en comparación con el de un imán común, y sin embargo lo bastante fuerte para orientar la aguja magnetizada de una brújula paralelamente a la dirección principal del campo. Éste converge en los polos magnéticos terrestres norte y sur, que no coinciden con los polos geográficos, sino que se hallan a cierta distancia de ellos. El campo de la Tierra no es estable; cambia día a día y durante largos lapsos que pueden ser años y siglos. Hay variaciones en la dirección de la brújula y en la intensidad del campo, como también perturbaciones conocidas como tormentas magnéticas. En unos pocos años, los polos magnéticos pueden alterar notablemente su posición.

Los minerales que reaccionan ante los campos magnéticos pueden retener las características que prevalecían en el campo terrestre cuando se formaron, o depositaron. Mediante ensayos con esos minerales y rocas, los geofísicos han podido obtener información sobre las antiguas condiciones geomagnéticas en todos los continentes. Esta clase de estudio se denomina *paleomagnetismo*, y se ha desarrollado principalmente desde la década de 1940.

La evidencia paleomagnética indica que, en el pasado, el campo terrestre era a menudo notablemente distinto del actual; por ejemplo, los polos magnéticos parecen haber ocupado lugares radicalmente diferentes en diversas épocas; o sea, que se habrían trasladado miles de kilómetros. Puesto que los polos geográficos de hoy en día nunca parecen distar mucho de los magnéticos, los paleomagnetólogos creen que los polos geográficos también se desplazaron junto con los magnéticos. Los geólogos antiderivistas han dudado de esta conclusión. En realidad, algunos niegan el valor de los estudios paleomagnéticos en general; pero se ha desarrollado una teoría sobre nomadismo polar –o cambios en la ubicación de los polos– junto con la teoría de la deriva continental. El nomadismo polar se basaba anteriormente en datos climáticos y fósiles. Ahora se ha visto reforzado por los descubrimientos paleomagnéticos y se lo ha empleado para respaldar la teoría de la deriva.

Empero, la interpretación del desplazamiento polar encierra una serie de ambigüedades. Tienen que haber ocurrido cambios de posición en los polos en relación con la corteza como un todo y con cada uno de los continentes. El eje terrestre en sí no se desplazó, como en un movimiento de todo el planeta. De hecho, los recorridos de los polos a través de los tiempos han sido tan complejos, que los geólogos han llegado a la conclusión de que, en realidad, ha sido la deriva variada de los continentes lo que causó esta distribución de compleja apariencia. Como esto implica el concepto de movimiento relativo, significa que no siempre se puede decir en forma definitiva si son los continentes o los polos los que se desplazaron, ni en qué direcciones lo hicieron.

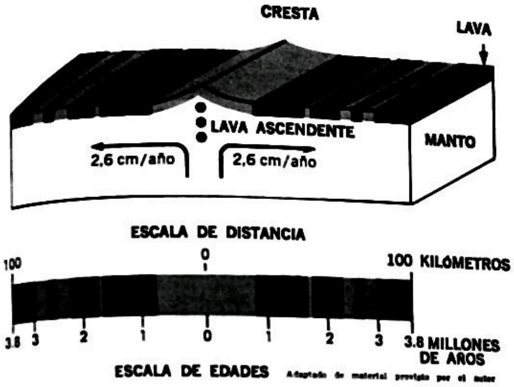
Los minerales paleomagnéticos también pueden indicarnos las latitudes geográficas que ocupaban los continentes en el momento en que dichos minerales se formaron. De nuevo, estas latitudes difieren de las que tienen actualmente los continentes, lo que podría explicar las extrañas condiciones climáticas de los tiempos prehistóricos, ya esbozadas en el trascurso del artículo.

El paleomagnetismo anotó otro triunfo importante para la teoría de la deriva. Se ha descubierto que el campo magnético terrestre puede invertir súbitamente su polaridad; el polo norte magnético se trasforma en el sur, y viceversa. En el pasado geológico, esta inversión del campo ocurrió muchas veces, pero, en apariencia, con mayor frecuencia en los últimos pocos millones de años. Estas inversiones quedaron impresas en los minerales magnéticamente sensibles, a tal punto que los geocronólogos están estableciendo una sucesión temporal mundial para los sucesos magnéticos, con el fin de datar rocas y fósiles. En la sección siguiente sobre geología marina señalaremos cómo esto ha permitido a los geólogos determinar e inclusive calcular la velocidad de desplazamiento y estiramiento de la corteza.

Geología marina. La geología marina –o submarina– es simplemente la aplicación de todos los campos de la geología al estudio de las costas, cuencas y fondos oceánicos. Hemos descrito la naturaleza de las grietas y los cordones mesoceánicos y su relación con el estiramiento cortical. Algunas de estas grietas llegan a los continentes en lugares tales como África oriental e Islandia, entre otros. También existe un gran sistema mundial de fracturas o fallas que cortan las grietas trasversalmente o casi en ángulo recto. Estas fallas se han relacionado con ciertas grandes fallas continentales a través de la Tierra, y se supone que las grietas mesoceánicas evidencian la separación de continentes unidos en otros tiempos.

Al datar las rocas y otros depósitos del fondo oceánico, los geólogos han descubierto dos hechos insólitos: ninguno tiene más de unos pocos cientos de millones de años de edad (en comparación con los varios miles de millones de años de ciertas rocas continentales); y cuanto más antiguas son las rocas, tanto más alejadas están de una grieta mesoceánica dada, a cada lado de la misma. Basándose en este hecho, han arribado a la conclusión –pese a las objeciones de los geólogos conservadores– de que las cuencas oceánicas son formaciones terrestres relativamente recientes. La datación peculiar de las rocas marinas muestra que la corteza se está apartando de cada lado de un cordón, y que se están formando nuevas rocas corticales en el emplazamiento de la grieta, para reemplazar a las antiguas. Esto explica por qué las rocas más jovenes se hallan cerca de la grieta y del cordón.

El estudio de las inversiones del campo magnético, descritas en la sección precedente,



Adaptado de material provisto por el autor

Expansión de la corteza a ambos lados de un cordón mesoceánico. A medida que el material cortical se aleja del cordón por obra de corrientes en el manto subyacente, surge más lava desde abajo para reemplazarlo. Las rocas más jóvenes están más próximas al cordón, y paralelas a él a cada lado. A medida que esto continúa, en las rocas nuevas se graban las variaciones del campo magnético terrestre. En las fajas de lava paralelas al cordón queda registrado un sistema de inversiones magnéticas alternadas, en el curso de varios millones de años. Las fajas oscuras representan cambios magnéticos negativos; las más claras, cambios positivos.

aporta más pruebas en favor de esta idea. Dado que el fondo oceánico se genera en el emplazamiento del cordón, y es arrastrado en ambas direcciones, debería estar compuesto por franjas alternadas de roca normal y opuestamente magnetizada, paralelas al propio cordón, a ambos lados del mismo. Los estudios hechos en el lugar han confirmado esta idea. El ancho de las franjas guarda la misma relación que la escala de tiempo de las inversiones del campo magnético terrestre y, a partir de ésta, se ha obtenido la rapidez de la expansión cortical, que puede llegar a 12 centímetros por año.

Basándose en estos datos, y otros más, los geólogos, con la ayuda de computadoras, han podido graficar las posiciones sucesivas de los continentes, a medida que derivaban en el trascurso del tiempo. Estudios recientes indican que, quizá, los continentes no derivaron individualmente, sino que varios segmentos mayores de la corteza terrestre pueden haberse desplazado. Cada segmento, o placa, contiene uno o dos continentes, junto con el fondo oceánico adyacente.

RESUMEN DE LA HISTORIA DE LA DERIVA

De este modo, muchos geólogos han reconstruido los movimientos de los continentes durante más o menos los últimos 500.000.000 de años de la historia de la Tierra. Alrededor de 250.000.000 de años atrás, toda la tierra firme del planeta estaba concentrada en dos masas continentales gigantescas: Laurasia, al norte, y Gonduana, al sur. Ambas estaban separadas por un extenso mar, pero, ¿cómo se formaron estos dos supercontinentes? Probablemente, de la fusión de otros más pequeños. Por ejemplo, en tiempos del Paleozoico Inferior (de 400 a 600 millones de años atrás) Laurasia se formó por la unión de secciones de Europa y Asia (Eurasia), por un lado, y de América del Norte por el otro. Eurasia y América del Norte estaban entonces separadas por el océano Protoatlántico (no el Atlántico actual) que desapareció al juntarse ambas masas continentales para formar finalmente Laurasia, hacia el Paleozoico Tardío. En la época siguiente –la era Mesozoica, o Edad de los Reptiles– Laurasia y Gonduana comenzaron a desmembrarse. El océano Atlántico empezó a abrirse nuevamente; comenzaron a formarse el océano Índico y el mar Antártico. A medida que Gonduana se desintegraba, América del Sur, África, Madagascar, la India, Australia, Antártida y, quizás, Nueva Zelandia, comenzaron a separarse. El proceso continuó durante la era siguiente, la Cenozoica, o Edad de los Mamíferos, hasta que se llegó a la geografía actual. Sin embargo, durante la deriva África y la India chocaron con Eurasia y quedaron unidas a ese continente, en especial la India. Se cree que esta colisión, ocurrida durante el Cenozoico, plegó los bordes de los continentes, formando los Alpes en Europa y los Himalaya en Asia.

No obstante, debemos recordar que todo esto puede haber sido simplemente la última fase de la deriva continental: ¿Quién sabe cuántas veces se han repetido tales sucesos en los períodos anteriores de la Tierra?

TERREMOTOS

Es difícil darse cuenta de que la corteza terrestre está cambiando constantemente. Sin embargo, hay fuerzas gigantescas que actúan sin cesar, moldeando una y otra vez las rocas corticales, empujándolas hacia arriba en forma de enormes pliegues, retorciéndolas y agrietándolas. Por lo general, este proceso es lento y raramente nos apercibimos de él, pero a veces se produce un gran terremoto, la tierra se sacude violentamente, en su superficie se forman largas grietas, se desmoronan monumentos y derrumban edificios con estrépito; cientos e incluso miles de personas mueren. Entonces comprendemos lo formidable de las fuerzas que actúan bajo la superficie de la Tierra.

Si bien las sacudidas destructivas son comparativamente pocas en cantidad, los terremotos se producen a menudo. Si incluimos en nuestro cómputo todas las sacudidas, desde las más grandes hasta las más pequeñas, cada año acontece como un millón de temblores. Además, hasta donde sabemos, han estado sucediendo en gran cantidad desde los tiempos más remotos de la historia de la Tierra. Por supuesto, los registros escritos de la humanidad se refieren frecuentemente a ellos.

Han existido muchas leyendas fantásticas sobre terremotos, pero apenas más fantásticas que las teorías sobre temblores propuestas por hombres cultos en el pasado. Sin embargo, en nuestros días, la investigación en el terreno mismo y el empleo de instrumentos de precisión de largo alcance han levantado gradualmente el velo de misterio. El estudio de los terremotos constituye ahora una ciencia aparte, conocida como *sismología*, de dos palabras griegas que significan "ciencia de los terremotos". Los científicos que se dedican a ella se denominan *sismólogos*.

CAUSAS

Los terremotos pueden estar causados por un mecanismo único, o por una combinación de ellos. Generalmente, a estas causas se las clasifica en cuatro categorías: (1) tectónicas, (2) volcánicas, (3) debidas al hombre, o artificiales, y (4) diversas.

MOVIMIENTOS TECTÓNICOS

La mayoría de los terremotos está causada por procesos tectónicos, movimientos a lo largo de *fallas*, esto es, fracturas en la corteza terrestre. En tales movimientos, un segmento de corteza, a un lado de la falla, experimenta un corrimiento respecto de otro, situado en el lado opuesto. Se cree que este proceso explica la mayor parte de los terremotos que ocurren cerca de la superficie de la Tierra.

Una teoría reciente sobre la corteza –conocida como teoría de la tectónica de placas, deriva continental, o expansión del lecho marino– puede proporcionar una explicación de esta clase de terremotos. Según ella, se están formando continuamente nuevas rocas en la corteza terrestre. El material rocoso surge desde el interior de la Tierra y emerge a lo largo de cordilleras del fondo oceánico. A medida que este material sube, hace que dichas cordilleras se separen, fenómeno conocido como expansión del lecho marino.

Muchos científicos creen también que la corteza terrestre se divide en varias placas que se mueven con relación a las otras. El material nuevo formado en las zonas de las cordilleras empuja la placas, apartándolas. A medida que las placas son apartadas en la zona de una cordillera mesoceánica, convergen en otras partes de la Tierra. Donde entran en contacto placas continentales y oceánicas, el borde anterior de la oceánica desaparece, o se hunde bajo la placa continental. Esta zona de eliminación origina buena parte de la actividad sísmica de gran intensidad y, por lo general, se la clasifica como de mucho riesgo en lo referente a terremotos. Cualquier zona de contacto entre placas –el punto en el que una debe ceder ante la otra– es una zona donde hay grandes probabilidades de que ocurran temblores.

ERUPCIONES VOLCÁNICAS

Algunos terremotos no son tectónicos, sino ocasionados por volcanes. Se deben a explosiones, o fracturas, que ocurren dentro de la estructura de un volcán. Cerca de éste pueden ser violentos, pero sus efectos no se perciben a distancias considerables.

Generalmente, los sismos volcánicos son de intensidad débil a moderada. Se cree se deben al *magma*, o roca fundida, que asciende a través de la corteza terrestre. Los terremotos de este tipo a menudo son precedidos por indicios de que está a punto de producirse una erupción volcánica.

TERREMOTOS OCASIONADOS POR EL HOMBRE

En forma totalmente accidental se descubrió que las actividades humanas podían iniciar terremotos de pequeña a mediana intensidad. En 1961, el ejército de los Estados Unidos usó un pozo profundo para deshacerse de materiales radiactivos; poco tiempo después, en las proximidades se sintieron temblores. En verdad, en un lapso de siete años se registraron mil seiscientos temblores, y se concluyó que había relación entre ellos y la inyección de líquido en el pozo. Se han registrado fenómenos similares cuando se inyectaron fluidos en campos petrolíferos con el fin de incrementar la producción.

También producen temblores de intensidad reducida a intermedia los ensayos nucleares subterráneos. Se desconoce, empero, si una gran explosión atómica podría desencadenar un terremoto de gran intensidad.

CAUSAS DIVERSAS

Varios otros factores pueden también tener influencia como causas de terremotos. Los estudios indican que, en algunos casos, existe una relación entre precipitación pluvial, agua subterránea y terremotos. También pueden tener influencia los choques de meteoritos contra la Tierra.

EL TEMBLOR EN SÍ

Las tensiones y distorsiones dentro de la estructura rocosa de la Tierra, cualesquiera sean sus causas, toman mucho tiempo para concretarse, pero al final, una ligera fuerza adicional puede ser la gota que rebase el vaso y hacer que se produzca una fractura en las rocas. Tal como hemos señalado, la mayoría de los terremotos parece ser causada por movimientos repentinos a lo largo de fallas. Por hallarse comprimidos unos contra otros, y como consecuencia de la fricción, estos bloques de roca no se desplazan libremente a lo largo de la superficie de la grieta; virtualmente quedan

© Phyllis McCutcheon Mithassel/PR

Foto que revela cómo un fuerte terremoto en Anchorage, Alaska, combó caminos pavimentados y derrumbó edificios. El temblor también se hizo sentir en las costas de California y en Hawai como olas de marea.

trabados entre sí, y cuando las fuerzas terrestres actúan, al principio no se mueven. En estas rocas flexionadas se generan grandes tensiones. Al alcanzar su límite de elasticidad –o punto de ruptura– se deslizan repentinamente paralelas a la superficie de la falla, precipitándose para ocupar su nueva posición, y la súbita sacudida ocasiona un terremoto.

El sitio de la estructura rocosa en que se produce el movimiento se denomina *foco* o *hipocentro*. El punto de la superficie terrestre que está directamente encima del foco se llama *epicentro*. Se han localizado focos de terremotos a diversas profundidades, desde cerca de la superficie hasta 800 y más kilómetros por debajo de ella.

El movimiento crea una perturbación comparable a la infinitamente más leve que se ocasiona cuando un guijarro cae en un estanque. Al producirse un estallido en la corteza terrestre, la perturbación asume la forma de ondas de amplitud creciente, que sin embargo difieren en varios aspectos de las ondulaciones en un estanque. Ante todo, efectúan su recorrido también a través del cuerpo de la Tierra, y no sólo en su superficie: además, viajan en dos grupos a través del planeta, y éstos se desplazan a velocidades diferentes.

ONDA DE COMPRESIÓN

La onda primera y más rápida se denomina *onda de compresión*, *onda primaria* u *onda P*. Proviene de variaciones en el volumen de las partículas terrestres al comprimirse y expandirse alternadamente en la fractura de la corteza. Estas compresiones y expansiones se trasmiten a través de la Tierra con una velocidad promedio de ocho kilómetros por segundo; la velocidad exacta depende de la naturaleza de los estratos a través de los cuales se desplaza la onda. Las partículas terrestres oscilan en la dirección de avance de la onda de compresión; en consecuencia, a menudo se llama a ésta *onda longitudinal*.

ONDA DE CORTE (O TRASVERSAL)

A la segunda onda sísmica se la denomina *onda de corte*, *onda secundaria* u *onda S*. La onda de corte de un terremoto se desplaza en todas direcciones; hace que las partículas terrestres vibren perpendicularmente a la dirección de movimiento de la onda; de ahí que se la llame, a veces, *onda trasversal*. Se desplaza a una velocidad promedio de cinco kilómetros por segundo. Tal como en la de compresión, el valor exacto de la velocidad depende del medio a través del cual avance.

EFECTOS

El movimiento impreso a la corteza terrestre por ondas sísmicas puede ser tan leve que nadie lo sienta. Nos enteramos únicamente porque lo registra cierto instrumento sensible llamado *sismógrafo*. En otros casos, los temblores pueden llegar a sentirse netamente y causar sacudidas y ruidos, pero no producirán cambios visibles en la superficie de la Tierra ni ocasionarán daño considerable. La gran mayoría de los terremotos corresponde a algunos de estos dos tipos.

Los terremotos graves –como éste en Nigata, Japón, en el que se abatieron casas de departamentos– ocasionan serios efectos posteriores: incendios, contaminación del agua y desprendimientos

NOAA Centro Nacional Geofísico y Solarterrestre de Datos

Empero, en otros temblores, la sacudida de la tierra es mucho más violenta. El suelo cimbra por todos lados. Si el terremoto es desusadamente fuerte, las casas y otras estructuras pueden ser desplazadas de sus cimientos o venirse abajo.

En los temblores serios pueden también producirse alteraciones notables de la superficie de la Tierra; formarse grietas en su corteza, y las secciones contiguas a ellas desplazarse horizontalmente hasta seis metros. La superficie puede también deslizarse verticalmente, y elevarse una de sus secciones por encima de otra adyacente. Asimismo, quizá se formen brechas en la tierra, pero en general son comparativamente poco profundas.

SONIDOS

Por lo común, los sonidos que acompañan a un terremoto inspiran pavor. Muchos de ellos no provienen de la Tierra, representan el violento resquebrajamiento de edificios y el ensordecedor estrépito de la caída de ladrillos, mampostería, revoque y objetos similares. No obstante, otros sonidos surgen realmente del suelo; las vibraciones producidas por la fractura de las rocas se trasmiten al oído humano, que puede captarlas únicamente si se hallan dentro de cierta escala de frecuencia. La menor frecuencia audible por el común de las personas es de unas veinte vibraciones por segundo; algunos terremotos producen vibraciones por debajo de las audibles. En los casos en los que pueden oírse, generalmente no están muy por encima del umbral humano de audición y, en consecuencia, su tono es bajo.

PEQUEÑOS TEMBLORES

A veces se sienten temblores preliminares, antes de una o varias sacudidas principales. Después de un fuerte terremoto, el terreno próximo al epicentro puede continuar siendo perturbado por una serie de temblores menores, conocidos como secundarios, o *réplicas*, que pueden proseguir durante meses, o años, después del principal.

Los efectos secundarios de los terremotos llegan a ser mortales; suelen estallar incendios causados por la rotura de cañerías y tuberías maestras de gas, o cortocircuitos. El sacudimiento del suelo puede ocasionar desprendimientos de tierra o rocas.

PERTURBACIÓN DE MASAS DE AGUA

Hasta ahora hemos estado hablando de temblores que se hacen sentir en tierra. También se producen en gran cantidad en las áreas oceánicas del planeta, las que constituyen unas tres cuartas partes de su superficie. El sismo que ocurre en el mar se denomina *maremoto*. A los que están a bordo de naves en la vecindad de dichas perturbaciones les parece que el buque estuviese batiendo contra algún obstáculo sumergido, como un arrecife. Por lo general, los oficiales de los barcos informan la latitud y longitud donde ocurrieron estos maremotos.

Los terremotos pueden tener efectos singulares sobre las masas de agua calma. En éstas pueden aparecer largas olas oscilantes conocidas como *seiches*, las que, frecuentemente, se producen en lagos y bahías. Las más espectaculares son las ondas enormes generadas en el mar por ciertos maremotos, si bien algunas se deben a temblores originados en tierra firme. Quizá, los deslizamientos del fondo submarino provocados por ciertos sismos terrestres perturban el mar. Otras olas oceánicas gigantescas se deben a desplazamientos del lecho oceánico, a lo largo de fallas. Las aguas parecen retroceder y luego se precipitan de vuelta con gran fuerza. Sin embargo, cuando estas olas han recorrido grandes distancias, se empequeñecen. Las provocadas por maremotos se conocen comúnmente como *olas de marea*, o *marejadas*; naturalmente, este término no es adecuado, pues en modo alguno se deben a la atracción lunar, o solar, que produce las mareas. Un nombre más preciso es el de *olas marinas sísmicas*. Los japoneses las llaman *tsunamis*, nombre ampliamente adoptado por los sismólogos.

Las tsunamis pueden recorrer grandes distancias a alta velocidad. El 1º de abril de 1946, las causadas por un maremoto frente a las islas Aleutianas llegaron a Valparaíso, Chile (a unos 12.000 kilómetros), en 18 horas, lo que indica que su velocidad era de alrededor de 660 kilómetros por hora.

No se pueden observar tsunamis en alta mar, sobre todo porque el volumen total de agua afectado es pequeño en comparación con el del océano. Además la distancia entre la cresta de una ola marina sísmica y la siguiente puede ser de hasta 160 kilómetros. Sin embargo, cuando una tsunami alcanza la costa, sus efectos

LISTA DE TERREMOTOS FAMOSOS

Fecha	Sitio	Descripción
1755	Lisboa, Portugal	Tres sacudidas arrasaron gran parte de la ciudad; una ola marina sísmica y el fuego coadyuvaron la destrucción. Más de 20.000 muertos.
1783	Calabria, Italia	Seis sacudidas principales en febrero y marzo. Muchos pueblos arrasados; 60.000 vidas perdidas. Cientos de sacudidas posteriores.
1811-12	New Madrid, Misurí	Serie de temblores violentos. Desprendimientos de tierra; áreas elevadas y deprimidas. Efectos percibidos hacia el este hasta Boston.
1886	Charleston, Carolina del Sur	El peor terremoto en la costa atlántica. Zona perturbada: casi 4.800.000 kilómetros cuadrados. Daños comparativamente reducidos en la mayor parte, pero Charleston sufrió pérdidas considerables.
1887	Sonora, México	El temblor elevó allí una parte de la cadena montañosa. Muchos pueblos y aldeas del valle destruidos.
1897	Assam, India	Las colinas de Assam ascendieron de cuatro a seis metros. Casi todos los edificios de mampostería de la región se derrumbaron; bosques destruidos, jungla y sembradíos inundados.
1905	Kangra, India	Temblor en las estribaciones de los Himalaya, en la provincia de Unjab. Se sintió en un área de 2.400.000 kilómetros cuadrados. 20.000 vidas perdidas.
1906	Formosa	Hubo desplazamientos verticales y horizontales. 1.300 muertos. Más de 7.000 edificios destruidos.
1906	California	Grieta de la corteza terrestre de unos 400 kilómetros. En San Francisco, el fuego infligió más daños que el temblor. La pérdida de vidas en todas las zonas afectadas totalizó alrededor de 1.000.
1908	Messina y Reggio, Italia	Ciudades parcialmente destruidas. Daños debidos principalmente a la deficiente construcción de los edificios y a la estrechez de las calles. Una ola marina sísmica coadyuvó la destrucción. Pérdidas de vidas estimadas entre 100.000 y 125.000.
1920	Kansu, China	Cientos de ciudades y pueblos destruidos; más de 200.000 muertos.
1923	Isla de Honshu, Japón	Tokio, Yokohama y otras ciudades afectadas por el temblor principal. Más de 140.000 muertos, más de 500.000 edificios destruidos.
1934	Bihar, India, y Reino de Nepal	Serie de desprendimientos de tierra y aludes en los Himalaya. Muchos pueblos destruidos o dañados. Se sintió en una zona de unos 322.000.000 de kilómetros cuadrados.
1939	Concepción, Chile	La ciudad de Concepción prácticamente destruida; otras veinte devastadas. Alrededor de 30.000 muertos.
1939	Turquía Central	Serie de temblores que sacudieron a casi toda Anatolia; ciudades, aldeas y pueblos destruidos. Número de muertos estimado en 50.000.
1946	República Dominicana	Uno de los mayores maremotos jamás registrados; grandes olas sísmicas. Secciones costeras devastadas. Grandes daños en Puerto Rico.
1949	Ecuador	Ambato y otras ciudades parcialmente destruidas. 6.000 muertos.
1950	Tibet, China, Birmania, India	Temblor resultante de un levantamiento en montañas del Tibet, en el extremo oriental de los Himalaya. Se sintió en una extensa zona de Tibet, China, Birmania e India. Geografía de la región considerablemente alterada.
1951	América Central	En El Salvador dos pueblos destruidos; graves daños también en Nicaragua.
1956	Afganistán del Norte	Muchas aldeas destruidas; más de 2.000 muertos.
1957	Irán	Norte y oeste de Irán reducido a ruinas; 2.900 muertos.
1960	Chile	Muchas aldeas destruidas; 5.000 muertos o desaparecidos.
1960	Marruecos	Dos temblores, incendios consiguientes y olas de marea mataron a 10.000 personas; Agadir en ruinas; 45.000 personas sin hogar.
1962	Noroeste de Irán	200 ciudades destruidas; 10.000 muertos.
1963	Skoplje, Yugoslavia	En ruinas la mayor parte de la ciudad; 1.000 muertos; 18.500 personas sin hogar.

LISTA DE TERREMOTOS FAMOSOS

Fecha	Sitio	Descripción
1964	Sur de Alaska	"El Terremoto de Viernes Santo" (27 de marzo): uno de los más graves conocidos; muchas poblaciones severamente dañadas; más de 100 muertos o desaparecidos; las olas de marea resultantes arrasaron costas de América del Norte, Siberia y Japón.
1966	Este de Turquía	60 aldeas en ruinas; 2.477 muertos; 100.000 personas sin hogar.
1970	Oeste de Turquía	La mayor parte de la ciudad de Gediz en ruinas, así como las aldeas circundantes. Más de 1.000 muertos; 90.000 personas sin hogar.
1970	Norte del Perú	Uno de los peores terremotos recientes; quizá 70.000 muertos.
1972	Sur de Irán	5.000 muertos; 45 aldeas destruidas.
1972	Nicaragua	La mayor parte de Managua en ruinas; más de 10.000 muertos.
1973	Sur de México	Ciudad de México y otras ciudades dañadas; cientos de muertos.
1974	Norte de Pakistán	Nueve pueblos fuertemente dañados; 5.000 muertos.
1975	Norte de China	Dañadas la ciudad de Yinkow y zonas aledañas. Cientos de muertos.
1975	Este de Turquía	Destruidos el pueblo de Lice y muchas aldeas; más de 2.000 muertos.
1976	Guatemala	Muy dañada Ciudad de Guatemala; la mayoría de los pueblos y aldeas de las zonas altas destruida; más de 23.000 muertos; más de 1.000.000 de personas (más de un quinto de la población) sin hogar.
1976	Noroeste de Italia	Casi 1.000 muertos; más de 50.000 personas sin hogar.
1976	Uzbiekistán, U.R.S.S.	Temblor muy fuerte; poca información disponible.
1976	Zona fronteriza entre China y Birmania	Dos temblores muy fuertes; poca información disponible.
1976	Irán Occidental, Nueva Guinea	Fuerte temblor que destruyó aldeas y mató a cientos de personas; posteriores desprendimientos ocasionaron casi 6.000 muertos.
1976	Noroeste de China	Dos temblores muy fuertes –de intensidades 8,2 y 7,9 en la escala de Richter– con 16 horas de diferencia; se cree que casi destruyeron la ciudad de Tangshan, matando o hiriendo a muchas personas.
1976	Mindanao, Filipinas	Fuerte temblor y olas de marea mataron a más de 3.000 personas en el área del golfo de Moro.
1977	Océano Índico	Uno de los más fuertes maremotos registrados; ocurrió frente a la isla indonesia de Sumba; enorme tsunamis dañadon aldeas.

UPI

Escena de una calle en Managua, Nicaragua, poco después de un violento terremoto.

suelen ser devastadores. Esto ocurre particularmente si la costa está en declive. Primero, el agua se retira a todo lo largo de ésta como si fuese una marea extraordinariamente baja; luego, una compacta muralla de agua vuelve e inunda la tierra. Pueblos enteros han sido destruidos por semejantes olas, que pueden alcanzar una altura de 15 metros, o más. Grandes buques anclados fueron barridos muy tierra adentro, donde quedaron varados.

GRADO DE TEMBLOR

Los terremotos se miden según su magnitud e intensidad. La magnitud es la cantidad de energía liberada por el temblor. El método más comúnmente empleado para medirla es la Escala de Magnitudes de Richter; se basa en la medición de la cantidad de movimiento del terreno, según lo determina un sismógrafo situado a distancias conocidas del epicentro del temblor.

La magnitud se expresa en números de uno a nueve. La escala es logarítmica: un incremento de un entero de la escala indica uno de diez veces en la magnitud del terremoto.

La intensidad de un terremoto es una medida de sus efectos. Se basa en la observación. Una escala de intensidades ampliamente utilizada es la de Mercalli.

El terremoto que sacudió a San Fernando, California, el 21 de junio de 1971 causó extensos daños a muchas estructuras subterráneas, tales como esta boca de tormenta.

Cuerpo de Ingenieros del Ejército de los E.U.

EL SISMÓGRAFO

Afortunadamente, los sismólogos han perfeccionado un ingenioso instrumento que puede detectar el movimiento de la corteza terrestre ocasionado por un terremoto, aun cuando ocurra a miles de kilómetros; este instrumento es el sismógrafo. Con este dispositivo es posible registrar las sacudidas y movimientos de los terremotos y determinar dónde se originan. También es posible obtener mucha información valiosa respecto a los estratos a través de los cuales pasan las ondas del temblor de tierra.

Un sismógrafo es fundamentalmente un péndulo. Se basa en el principio de que, a causa de su inercia, o resistencia a los cambios de movimiento, la masa pesada del disco del péndulo que está en su extremo permanecerá inmóvil cuando se agita el suelo que tiene debajo. Una pluma registradora solidaria con el disco registra este temblor en un papel que se mueve junto con el terreno.

El papel para registro está enrollado en torno de un tambor que se mantiene girando y moviéndose también hacia adelante, bajo la pluma: como resultado, ésta traza una línea continua en el papel.

En realidad, para evitar la fricción, la "pluma" registradora generalmente es un punto luminoso que se refleja desde un espejo, produciendo una delgada línea negra sobre un papel fotográfico. Los casi imperceptibles valles y crestas causados por las vibraciones se aumentan grandemente mediante un proceso de ampliación. Un dispositivo señala las horas y minutos de cada registro.

Para registrar adecuadamente el movimiento del suelo ocasionado por un temblor se necesitan tres sismógrafos: uno vertical y dos horizontales. Si se obtienen registros en los tres puede calcularse la distancia desde el observatorio hasta el epicentro.

LOCALIZACIÓN DEL EPICENTRO

El tiempo de llegada de la onda de compresión, o P, se lee en la hoja de registro del sismógrafo vertical. El de la onda de corte, o S, se lee en forma similar en uno u otro de los horizontales. Se determina el intervalo que media entre la llegada de la onda P y la de la S. La distancia del sismo puede determinarse entonces por medio de tablas. Supóngase, por ejemplo, que la onda S llega a un observatorio dado 6 minutos y 40 segundos después que la P; consultando las tablas, vemos que la distancia del arco entre observatorio y epicentro del terremoto es de 4.900 kilómetros. La distancia de arco entre dos puntos de la Tierra corresponde a la línea más corta posible que conecta los puntos y sigue la curvatura del planeta.

Pero aun si conocemos la distancia entre nuestro observatorio y el epicentro de un terremoto dado, nuestra tarea no está completa. Sólo sabemos que el sismo ocurre en algún lugar de la circunferencia de un círculo, con el observatorio en su centro y la distancia desde éste al temblor como radio. Para descubrir dónde se halla el sismo nos ponemos en contacto con otros dos observatorios situados en localidades bien separadas, y que también hayan registrado el terremoto. Construimos tres círculos en un mapa, centrados en cada observatorio, y con radio igual a sus distancias desde el epicentro del sismo; el punto donde se intersecan es el epicentro.

La destrucción de carreteras públicas por los terremotos es tan costosa como inconveniente. También puede ser mortal, particularmente si ocurre durante las horas de máximo tránsito, cuando las calles están atestadas.

PREDICCIÓN

En la década de 1970 los científicos comenzaron a perfeccionar métodos de predicción de terremotos. Uno se basa en un efecto medible llamado dilatancia, que ocurre en rocas someti-

Escala de intensidad de terremotos. Esta escala indica la cantidad de daño superficial causado por un terremoto (Escala Modificada de Intensidades de Mercalli, de 1931).

I. Muy pocos lo perciben, bajo circunstancias especialmente favorables.
II. Sólo lo perciben unas pocas personas en reposo, especialmente en pisos altos.
III. Percibido con bastante nitidez en interiores, sobre todo en pisos altos, pero muchos no lo reconocen como temblor. Los autos parados pueden oscilar levemente. Vibración como la de un camión.
IV. De día, muchos lo notan en interiores; afuera, pocos. De noche, algunos despiertan. Vibran platos, ventanas, puertas; las paredes crujen. Sensación como de un camión golpeando el edificio. Los autos parados oscilan perceptiblemente.
V. Notado por casi todos; muchos despiertan. Hay platos, ventanas, etc., rotos. Pocos casos de revoque agrietado. Objetos inestables volcados. A veces, perturbación en plantas, postes y otros objetos altos. Los relojes de péndulo paran.
VI. Todos lo perciben; muchos se espantan y corren afuera. Se mueven algunos muebles pesados; pocos casos de revoques caídos o chimeneas dañadas. Escasos daños.
VII. Todos corren afuera; daños escasos en edificios de buen diseño y construcción; de leves a moderados en estructuras comunes bien hechas; considerable en las mal erigidas o diseñadas; chimeneas rotas. Advertido por choferes.
VIII. Ligeros daños en estructuras especialmente diseñadas; considerables en edificios sólidos, con derrumbe parcial; grandes en estructuras mal hechas. Tabiques derribados de las armazones. Caída de chimeneas, columnas, monumentos, paredes, muebles pesados volcados. Arena y lodo eyectados en pequeñas cantidades. Cambios del agua de pozo. Conductores de autos perturbados.
IX. Daño considerable en estructuras diseñadas especialmente; armazones bien proyectadas fuera de plomada; daños grandes en edificios sólidos, con derrumbe parcial. Edificios desplazados de sus cimientos. Suelo notablemente agrietado.
X. Estructuras de madera sólidas destruidas y la mayoría de las de ladrillo y armazón, junto con los cimientos; suelo muy agrietado; rieles torcidos. Deslizamientos de tierra en márgenes de ríos y pendientes abruptas. Arena y lodo movidos. Agua salpicada sobre las márgenes.
XI. Pocas o ninguna estructura de ladrillo en pie. Puentes destruidos. Anchas fisuras en el suelo. Tuberías subterráneas completamente inutilizadas. Derrumbes y deslizamientos de tierra en suelos flojos. Rieles retorcidos por la fuerza del sismo.
XII. Daño total. Las ondas se ven en la superficie del suelo. Distorsión de visual y de nivel. Objetos arrojados a los aires.

A veces, los terremotos causan una pequeña brecha en la superficie de la tierra, tal como revela la foto de arriba, que muestra los efectos de un fuerte temblor en el Perú. La foto de abajo muestra la falla de San Andrés, en California, zona en la que podría producirse pronto actividad sísmica, según algunos científicos.

das a presión poco antes de que las despedace un terremoto. La palabra dilatancia se refiere a la dilatación, o ensanchamiento, de las grietas de dichas rocas.

A fines de 1960, científicos de la Unión Soviética comenzaron a advertir que aparecía regularmente en sus sismógrafos una clase insólita de vibración poco antes de los terremotos. La vibración resultó ser producida por una anomalía en la velocidad de una clase de onda sísmica: la onda de presión. Ésta se retardaba, conservaba esa velocidad más baja durante un tiempo y luego volvía a la normalidad. Poco después se producía un terremoto. Varios científicos propusieron la teoría de dilatancia-difusión para explicar estos hechos.

Según esta teoría, a medida que crece la tensión sobre las rocas de una zona de falla, se desarrolla en ellas una red de grietas diminutas que se llenan con agua subterránea. Si la presión que actúa sobre las rocas comienza a crecer con gran velocidad, las grietas empiezan a dilatarse, o expandirse, y el agua subterránea drena de ellas. Las ondas de presión avanzan más lentamente a través del espacio seco que a través del que está lleno de líquido y, en consecuencia, el sismógrafo registra un frenamiento; pero, ulteriormente, el agua se infiltra de nuevo en las grietas expandidas, y la onda de presión vuelve a la normalidad. En este punto, las rocas, debilitadas por el amplio agrietamiento y por la presión renovada de agua, estallan, produciendo un terremoto.

Estos cambios de presión indicadores pueden medirse en las estaciones de predicción de sismos, las que han aumentado su número en China, Estados Unidos y en otras partes. Probablemente, también sea resultado de la dilatancia una combadura o abultamiento gradual de la superficie terrestre que se observa antes de los terremotos.

Otros signos de un próximo temblor incluyen variaciones en el nivel del agua en pozos profundos y un contenido desusadamente alto de gas radón radiactivo en el agua de pozo.

Los científicos han especulado también sobre algunas posibilidades desusadas de predicción. Por ejemplo, los sismólogos chinos han usado durante mucho tiempo en sus predicciones las observaciones de la conducta animal. Por razones desconocidas, algunos animales parecen mostrar tipos de comportamiento insólitos an-

NASA, U.S.I.S.

Vista de la ciudad de Managua, Nicaragua, mostrando la muy amplia zona de destrucción ocasionada por el temblor de 1972, que causó unas 10.000 muertes.

tes de que se produzca un temblor. Sin embargo, está lejos de establecerse la validez y confiabilidad de tales pautas de comportamiento.

Todavía más insólitos son los informes sobre el fenómeno de la "luz del terremoto". Desde hace mucho se han descrito extraños fogonazos que aparecen en el cielo poco antes de un gran sismo. Hasta hace poco, la mayoría de los científicos desestimaba dichos informes, pero ahora existen datos dignos de confianza y fotografías de estos fenómenos. Una explicación posible es que se genera un potencial eléctrico en ciertas rocas cuarcíferas cuando se las somete a una tensión tal como la que sufrirían a lo largo de una falla que estuviese a punto de ceder. El potencial podría liberarse como descarga eléctrica, iluminando el cielo. Aún no se ha investigado adecuadamente esta posibilidad.

MINERALES

Por lo general, la gente tiene una vaga idea de que los minerales están relacionados con las rocas. Recuerda, también, que la palabra "mineral" se usa para designar lo que no tiene vida y nunca la tuvo (como al hablar del "reino mineral"). En realidad, los minerales son los bloques que constituyen los cimientos de la Tierra y su corteza. Son las sustancias que forman las rocas.

Podemos definir un mineral como un compuesto o elemento químico inorgánico (sin vida) de origen natural. Por "origen natural" significamos hallado en el terreno, las rocas o aguas de la Tierra y no hecho por el hombre. Un mineral en su forma natural, sin explotar, es una *mena*. Como es una sustancia inorgánica, los seres vivientes y sus productos orgánicos inmediatos (como la miel de abejas) no pueden ser minerales.

Hay algunas excepciones, sin embargo. En ciertos casos, los restos de seres que vivieron antiguamente pueden experimentar cambios radicales y, con el trascurso del tiempo, llegar a convertirse en minerales. Una cantidad considerable de grafito mineral, una forma blanda de carbón hallado en las rocas, probablemente derivó del carbono en los tejidos de organismos muertos hace mucho tiempo y enterrados en limos que luego se convirtieron en roca. Ciertos organismos forman calcio, azufre o compuestos de hierro, que pueden quedar como depósitos minerales.

TIPOS DE MINERALES

Hay muchos tipos diferentes de minerales. Hasta el presente se conocen unas 1.500 variedades. De ellas, relativamente pocas –de 150 a 200– son comunes o están ampliamente difundidas. El número de minerales conocidos puede aumentarse considerablemente si contamos variaciones menores causadas por pequeñas diferencias en la composición química y la presencia de diversas impurezas. Los minerales pueden clasificarse en ocho grupos según los elementos de que se componen o de los compuestos que forman.

(1) Minerales como elementos. Estos minerales son sustancias simples que no están químicamente combinadas con otros elementos, sino que existen en estado libre. En algunos casos, esto se debe a que no son muy activos químicamente. También puede ser porque los elementos con los que normalmente se combinan han faltado, o porque las condiciones fueron desfavorables para que los compuestos se formaran o se conservaran. Buenos ejemplos de elementos naturales los constituyen el grafito y el diamante, que son diferentes formas del carbono. El azufre a menudo no está combinado; lo mismo algunos metales, incluyendo los tan conocidos como oro, cobre, plata y platino.

(2) Sulfuros. Son compuestos de metales con el elemento azufre. Las sulfo-sales generalmente se incluyen en este grupo. Los minerales raros denominados telururos y seleniuros son similares a los sulfuros en varios aspectos, pues el telurio y el selenio son parientes químicos del azufre.

Los sulfuros tienen ciertas propiedades físicas comunes. Muchos de ellos poseen brillo metálico, y en su mayoría son muy pesados.

La pirita de hierro es un buen ejemplo de un sulfuro brillante y pesado. Se parece notablemente al oro y algunas veces se la llama "oro del tonto". Las menas de muchos metales importantes pertenecen a este grupo. Entre ellas están la galena, la mena principal del plomo; el cinabrio, una mena del mercurio; y la blenda, o esfalerita, menas del cinc.

En varios casos, dos o tres sulfuros están constituidos por los mismos elementos químicos, pero con diferentes porcentajes de los elementos que los componen. En dichos casos, los nombres que llevan los minerales sirven para distinguir uno de otro (u otros). Por ejemplo, oropimente (As_2S_3) y rejalgar (AsS), ambos son sulfuros de arsénico; calcopirita ($CuFeS_2$) y bornita (Cu_5FeS_4), son ambos sulfuros de cobre y hierro.

(3) Óxidos. Son compuestos de varios metales con oxígeno. Van desde compuestos simples a combinaciones complejas de varios metales, oxígeno, hidrógeno y agua. Como los sulfuros, los óxidos pueden mostrar brillo metálico y ser pesados. Un óxido de aluminio, el corindón, es tan duro que se usa como abrasivo. Otros minerales en este grupo, como cincita (óxido de cinc) y cuprita (óxido de cobre), son muy blandos. Entre los óxidos se incluyen las menas importantes de cobre (cuprita), hierro (hematita), manganeso (pirolusita), etc.

(4) Sulfatos. Son compuestos de azufre, oxígeno y diferentes metales. Los sulfatos que contienen moléculas de agua y son solubles en ésta se conocen como sulfatos hidratados o hidrosulfatos. Las formas anhidras no contienen moléculas de agua y son mucho menos solubles en ésta que los sulfatos hidratados. El sulfato de calcio hidratado, que se conoce como yeso, y se emplea para el yeso París, es tan blando que puede rayarse con la uña. Los minerales bastante raros conocidos como cromatos, que incluyen en sus moléculas los elementos cromo y oxígeno, se agrupan a veces con los sulfatos.

(5) Carbonatos. Son compuestos de carbono y oxígeno con varios metales. Los carbonatos son blandos y sumamente solubles en ácidos. El

Katherine H. Jensen

Derecha: la wulfenita, un óxido de plomo y molibdeno, se presenta a menudo como cristales tabulares o columnares. Página opuesta: azurita (azul) y malaquita (verde), dos carbonatos minerales de cobre.

Katherine H. Jensen

El aguamarina es una variedad de piedra preciosa del mineral berilo. Los cristales son de características grandes y hexagonales y tienen un tinte azul o verde.

mineral más importante en este grupo es la calcita, o carbonato de calcio, que a menudo forma enormes masas de piedra caliza. Otro importante constituyente de rocas es la dolomita, o carbonato de calcio y magnesio. A veces, clasificados junto con los carbonatos hay minerales como los boratos (compuestos de distintos metales con boro y oxígeno) y los nitratos (compuestos de metales con nitrógeno y oxígeno).

(6) Haluros. Son sales en las que los metales están combinados con uno de los halógenos, un grupo de elementos que incluye flúor, cloro, bromo, yodo y astato. La sal común, o cloruro de sodio, es el ejemplo más conocido. Los depósitos de haluro se formaron a menudo por evaporación del agua –particularmente la de mar– en la que habían estado en solución. Como muchos haluros se disuelven fácilmente en agua, rara vez se encuentran como minerales, excepto en climas secos o donde han estado protegidos de filtraciones por rocas impermeables. Comúnmente se los encuentra junto con el carbonato de calcio y varios sulfatos.

(7) Fosfatos, tungstenatos, molibdenatos, uranatos, vanadiatos, arseniatos y otros. Los minerales representativos de este grupo constan de oxígeno más metaloides y metales tales como fósforo, tungsteno, molibdeno, uranio, vanadio o arsénico. Un número importante de menas se incluyen aquí. Por ejemplo, la wolframita es una fuente importante de tungsteno; la vanadita, de metal vanadio.

(8) Silicatos. Son compuestos de diferentes elementos con silicio y oxígeno. Constituyen una proporción considerable de las especies minerales y son los más importante formadores de rocas. Van desde compuestos simples a estructuras moleculares y cristalinas muy complejas. Empero, todos tienen en común la combinación de átomos silicio-oxígeno.

El miembro más simple del grupo es la sílice, o cuarzo, que contiene sólo átomos de silicio y oxígeno en la proporción de uno a dos. Es un mineral muy común, que se presenta en variedad de formas, y es el principal constituyente de la mayoría de las arenas. A menudo se lo encuentra disperso en las masas rocosas. Los feldespatos forman un gran grupo de silicatos metálicos que se hallan en muchos tipos de rocas. Otros silicatos comprenden las micas, las arcillas, talco y amianto. Ciertos silicatos se clasifican como gemas: entre éstas se hallan el granate y el circón.

FORMACIÓN DE MINERALES

Los minerales se han formado de distintas maneras. Ciertos tipos derivan directamente del magma, o roca fundida, que se solidifica formando roca ígnea. El cuarzo, las micas y los feldespatos son buenos ejemplos de rocas formadas de esta manera. Otros minerales, como la sal y el azufre, pueden depositarse como sólidos directamente del estado gaseoso, alrededor de las bocas volcánicas. La wolframita, el oro, el cuarzo y otros pueden precipitar directamente de una solución en aguas magmáticas calientes: vapor y agua que surgen de la roca líquida.

El tipo de depósito mineral denominado ***vena*** se ha formado a medida que las aguas circulantes han dejado caer sus cargas de materia mineral disuelta en una grieta de la roca. Entre los minerales de vena comunes se encuentran el oro, el cuarzo y los sulfuros. Ocasionalmente, el agua cargada de minerales revestirá un espacio hueco y redondeado en las piedras, con material cristalino. Un pedrejón hueco de materia mineral, recubierto interiormente de cristales de este tipo, se denomina ***geoda***.

Algunos minerales estaban originalmente en solución en masas de agua y formaron depósitos a medida que ésta se evaporaba. Entre los más comunes que se depositaron en tal forma se cuentan el yeso, otros sulfatos y la sal gema. Los carbonatos, boratos y nitratos también se formaron por el proceso de evaporación. Donde los depósitos son particularmente extensos, puede presumirse que en esas regiones prevalecieron antaño condiciones de aridez.

Ciertos minerales se desarrollaron de la alteración de otros preexistentes que han estado expuestos al desgaste por la acción de la intemperie. El oxígeno contenido en el aire y el agua ataca a menudo a los minerales en las rocas, convirtiéndolos en diferentes compuestos tales como óxidos y sulfatos. Algunos de estos productos son más solubles que los materiales originales y por tanto son arrastrados paulatinamente por las corrientes. En algunos casos, los minerales son mecánicamente fragmentados en trozos más pequeños, que luego el viento y el agua llevan y desintegran. Estos fragmentos se acumulan después y pueden ser cementados, formando roca sedimentaria. Cuando esas masas cementadas o, en realidad, cualquier clase de rocas, quedan expuestas a la acción del tiempo y la erosión, sólo las sustancias originales más durables e insolubles pueden subsistir y formar depósitos tales como sílice, arcilla y bauxita (la mena del aluminio). Nuevos minerales pueden surgir de la trasformación operada en las rocas por el calor, la presión, o ambos. Estas trasformaciones se conocen como metamorfismo. La calcita cristalina se forma cuando la piedra caliza sufre metamorfismo. El silicato olivina se trasforma en serpentina por acción de agua caliente derivada del magma. Los minerales metamórficos pueden ser alargados en forma de aguja, de láminas, o de placas. A veces son de tamaño y belleza sorprendentes. Las formas y posiciones que presentan en las rocas revelan frecuentemente la intensidad y direcciones de las fuerzas que han actuado sobre ellos.

Un mineral, entonces, puede considerarse como el resultado neto de la acción de fuerzas naturales sobre la materia terrestre en un área dada y durante un determinado lapso. Tomando como base el tiempo geológico, pocas formas y combinaciones minerales persisten mucho tiempo. Un mineral se convierte en otro o reacciona con otro a medida que lo afectan diferentes procesos.

Los geólogos cartografían la distribución superficial de los minerales. También estudian muestras de perforaciones hechas a varios kilómetros de profundidad en la corteza terrestre. Estas investigaciones pueden servir para fines prácticos, pero aun cuando no sea así, contribuyen a resolver el misterio del origen y desarrollo de la Tierra.

Institución Wards de Ciencias Naturales, Inc.

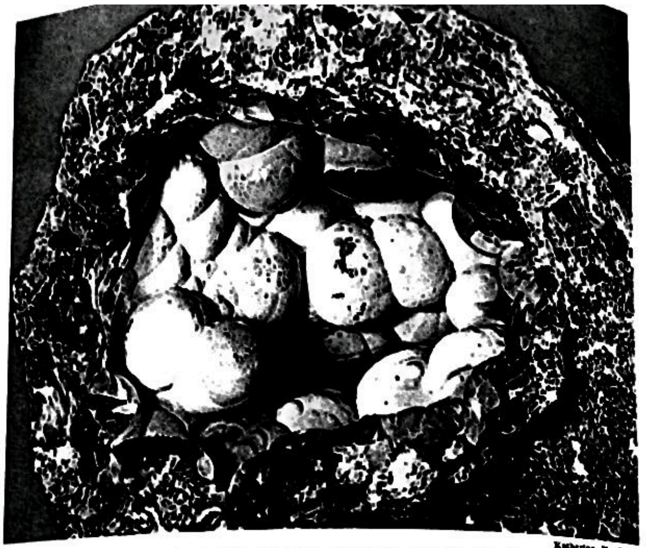
La piedra preciosa lazulita o lapislázuli, incrustada en mármol. El mineral es un compuesto complejo que contiene sodio, aluminio, sílice y azufre.

ESTRUCTURAS CRISTALINAS

Los minerales y otros sólidos verdaderos, naturales o sintéticos, tienden a formar cristales. Éstos son masas simétricas con definidas formas geométricas angulares. Sabemos ahora que las caras angulares externas de un cristal se relacionan con su estructura atómica interna. Los átomos o iones (átomos eléctricamente cargados) del cristal se disponen en posiciones relativas definidas, formando un esquema básico denominado ***retículo***. Esto se probó definitivamente en el año 1912, cuando el físico alemán Max von Laue utilizó los rayos X para el estudio de los cristales. Los rayos X han probado la existencia del retículo del cristal fuera de toda duda y arrojado mucha luz sobre la estructura cristalina. Cuando los rayos pasan a través de un cristal son dispersados, o difractados, por capas o planos atómicos dentro del cristal. Una imagen en "sombra" de los planos se proyecta sobre una película fotográfica que luego se revela. La estructura cristalina puede entonces deducirse del análisis del esquema de sombra. Los físicos británicos sir William H. Bragg y sir William L. Bragg (que eran padre e hijo) realizaron importantes investigaciones en el campo de la estructura cristalina.

Mucho antes que los rayos X revelaran la verdadera naturaleza de los cristales, éstos fueron profundamente estudiados y clasificados basándose en sus formas externas. Esta clasificación se usa aún, pues los descubrimientos de los rayos X no la contradijeron, y es adecuada para el estudio externo de los minerales.

Hay treinta y dos clases, en total, en esta



Katherine H. Jones

Una geoda de calcedonia. El mineral, una forma cristalina muy fina del cuarzo, reviste la cavidad de la roca, o geoda, con masas blancas aterronadas.

clasificación, y están agrupadas en seis divisiones principales, denominadas *sistemas*. Los sistemas se basan en el ordenamiento de tres (o de cuatro) ejes imaginarios que se intersecan en un punto dentro del cristal. En algunos casos, los ejes son de la misma longitud; en otros, son diferentes. Los ejes pueden ser perpendiculares entre sí o intersecarse en ángulos diversos.

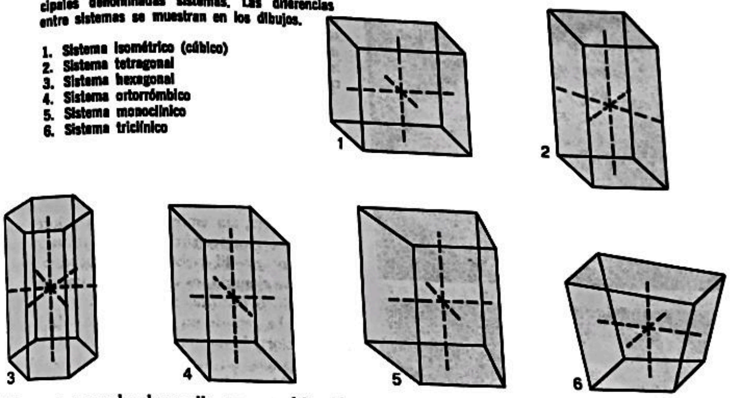
El *carácter del cristal*, o las formas individuales de los cristales y los modos de constituirse son una característica importante de los minerales que ayuda a identificarlos. Difícilmente un mineral alcanza una forma perfecta. Por lo general está distorsionado o quebrado de algún modo. Esto a menudo hace difícil determinar el sistema y clase a que pertenece. Además, el cristal puede adoptar diferentes formas aun dentro de un mismo sistema. Por ejemplo, la fluorita –fluoruro de calcio– forma cristales cúbicos cuando cristaliza a temperaturas bajas. A temperaturas mayores, sin embargo, se desarrolla en cristales de ocho caras (octaédricos).

A veces, un mineral como la calcita crece lentamente, adoptando diferentes formas en distintos momentos. La final puede ser muy diferente de las primitivas. Los cristales de calcita revelan a menudo fantasmas borrosos de etapas de crecimiento precedentes dentro de las capas trasparentes externas, que fueron las últimas en formarse. Se ve con claridad que en algunos casos la forma primitiva fue muy diferente de la que se desarrolló después. No obstante, no importa cuál sea el estado final de un cristal, sus direcciones internas están siempre referidas a los mismos ejes, de igual modo que los ángulos entre los mismos conjuntos de caras son siempre constantes.

Una muestra mineral puede encontrarse como un solo cristal grande, a veces de uno o dos metros de longitud, o como un cristal pequeño, a veces microscópico. Las masas de cristales o sus fragmentos pueden agregarse, unirse o esparcirse a través de una sección de roca. Un cuerpo cristalino puede tener casi cualquier forma concebible: de hoja, fibrosa, de aguja, de racimo, etc.; ello depende del mineral y de su carácter. Cristales de diferente composición suelen hallarse entremezclados. Otro fenómeno notable es la duplicación: dos cristales de un mineral pueden encontrarse unidos de distintos modos, a veces como gemelos siameses. El mineral estaurolita, un silicato de aluminio y

Los cristales se agrupan en seis divisiones principales denominadas sistemas. Las diferencias entre sistemas se muestran en los dibujos.

1. Sistema isométrico (cúbico)
2. Sistema tetragonal
3. Sistema hexagonal
4. Sistema ortorrómbico
5. Sistema monoclínico
6. Sistema triclínico



hierro, a menudo desarrolla una combinación cruciforme de dos cristales. Estaurolita es un nombre apropiado porque la palabra *stauros* en griego significa "cruz".

PROPIEDADES FISICAS

El análisis de las formaciones de cristales ayuda a identificar los minerales, pero otras propiedades físicas pueden ser aún más útiles para su identificación. Entre esas propiedades figuran dureza, color, veta, brillo, densidad, clivaje, partición, fractura y tenacidad.

Dureza. La dureza de un mineral a menudo da un indicio de su identidad. Depende del espacio entre las diversas capas atómicas del cristal. Las diferencias en el espaciado se traducen en diferencias en dureza, aun en una misma muestra de mineral. Por ejemplo, la calcita puede rayarse fácilmente con la uña a lo largo de la base de un cristal. No puede, sin embargo, rayarse cuando uno desliza la uña a lo largo de la cara por la que a menudo se rompe: la *cara de clivaje*.

Se han ideado varias escalas de dureza. La propuesta en 1822 por el mineralogista alemán Friedrich Mohs se basaba en la dureza relativa de diferentes minerales. Probó ser tan práctica y fácil de aplicar que aún se utiliza.

En la escala de Mohs, a ciertos minerales típicos se les asigna una dureza que va de uno a diez, como sigue:

1. Talco
2. Yeso
3. Calcita
4. Flúor
5. Apatita
6. Feldespato
7. Cuarzo
8. Topacio
9. Corindón
10. Diamante

El talco es el más blando de estos minerales; el diamante, el más duro. Un mineral dado, en

Estas varillas color de rosa son cristales de turmalina, un mineral que existe en todo el mundo. Los cristales son hexagonales y a menudo bastante grandes.

Izquierda, arriba: una muestra de ágata listada, rodeada de capas de cuarzo y amatista. El ágata listada es una forma multicolor de calcedonia. Abajo: amatista, una variedad trasparente del cuarzo púrpura o azul-violeta. A menudo se lo ve recubriendo la pared interna de una roca ahuecada. Este mineral es también una piedra preciosa muy común en joyería.

esta escala puede rayar cualquier mineral con un número más bajo y puede ser rayado por cualquier otro con un número más alto. Para determinar la dureza de un mineral que no está representado en la escala, probamos a rayar el mineral con uno de los minerales típicos clasificados antes. Comenzamos con el más blando y luego avanzamos en la escala. Si el mineral desconocido puede rayar la calcita, pero es rayado por la fluorita, su dureza está dada como 3½. Si se acerca más a la calcita que a la fluorita, su dureza podría darse como 3¼.

Se puede determinar la dureza de los minerales con mayor exactitud que con la escala de Mohs. Para la mayoría de los mineralogistas, sin embargo, las mediciones realizadas con la escala de Mohs son lo bastante precisas para fines prácticos.

Color. El color puede también ayudar a identificar distintos minerales. Empero no es siempre una guía segura por su naturaleza cambiante. Las impurezas en una muestra mineral determinan a menudo su tinte particular. Ciertos minerales, sin embargo, tienen colores sorprendentes y característicos y se han usado para hacer pinturas. El carbonato de cobre, o malaquita, por ejemplo, es de un hermoso verde. El lapislázuli, o lazurita, es ultramarino.

Veta. La veta de un mineral está relacionada con el color y la dureza. Un mineral más blando que, digamos, un azulejo de porcelana blanca (dureza 7) puede dejar una veta coloreada sobre éste, al ser raspado contra él. Aunque parezca extraño, el color de la veta es a menudo un medio más constante y certero de identificación que el color del mineral que produce la veta. Por ejemplo, muestras de hematita y limonita, ambos diferentes óxidos de hierro, pueden tener igual color externo: negro. Sin embargo, la hematita puede distinguirse de la limonita por su veta roja. La veta de la limonita es amarillo ocre o herrumbre.

Brillo. La propiedad denominada brillo, o lustre, es útil a menudo para identificar minerales. Un mineral puede tener un brillo metálico, grasoso, vidrioso, resinoso, sedoso, perlado o adamantino (como el diamante). Esta propiedad está íntimamente relacionada con la estructura interna del cristal y su carácter, que afecta la forma en que refleja o trasmite la luz. Algunos minerales no tienen realmente nada de brillo, pero pueden formar masas terrosas, opacas.

Densidad. La densidad es otra característica distintiva de los minerales. La densidad, o peso por unidad de volumen, de una sustancia depende del peso de los elementos que la forman y de cómo se encuentran concentrados en la estructura cristalina. La densidad se da a menudo como peso específico, esto es como la relación del peso de la sustancia en el aire con el peso de un volumen igual de agua. El peso específico del diamante es 3,5, mientras que el del grafito es 2,3. Esto puede parecer sorprendente, pues ambos son carbono. La explicación es que el diamante y el grafito difieren en estructura cristalina.

A menudo, minerales aparentemente similares, tales como la casiterita oscura y turmalina negra, pueden distinguirse fácilmente por diferencias notables en el peso específico, pues el de la casiterita es mucho mayor. Una vez que hemos adquirido experiencia, podemos distinguir aun las menores diferencias de peso entre volúmenes iguales de minerales, simplemente levantándolos.

Clivaje. El clivaje, o separación, se determina por la disposición atómica en un cristal. Si los átomos están densamente concentrados o fuertemente unidos, el clivaje es difícil. Es fácil dondequiera que los átomos estén ampliamente espaciados o débilmente ligados. Un mineral puede dividirse a lo largo de lo que se denomina ***plano de clivaje***, que es a menudo paralelo a una cara del cristal o una posible cara de éste. Puede haber varias direcciones de clivaje en un mismo mineral. En algunas de esas direcciones, el clivaje resulta más fácil que en otras. Un cristal cúbico de fluorita tiene clivaje octaédrico; esto significa que todas las esquinas del cubo pueden separarse completamente, resultando un octaedro, o cuerpo de ocho caras, que recuerda dos pirámides de cuatro caras dispuestas base contra base. Como el clivaje está relacionado con la estructura cristalina, es útil como una propiedad identificadora de los cristales.

Partición. La propiedad de la partición incluye división en determinado plano. Sin embargo, la partición no es la misma cosa que el clivaje, pues no se debe a la disposición normal de los átomos, sino a los ocasionales debilitamientos que pueden desarrollarse en una masa mineral por muchas razones químicas o físicas.

Fractura. La fractura, o ruptura, no está relacionada con ningún plano o dirección interna. Sin embargo, los minerales tienen tendencia a romperse en cierta forma. Algunos, especialmente los metales nativos, tienen una fractura serrada. El cuarzo, como el vidrio, da una ruptura concoidea, o sea que ésta exhibe una superficie curva, como un caparazón.

Tenacidad. La tenacidad representa la resistencia de un mineral a la ruptura bajo una fuerza o tensión. Un mineral quebradizo, como el cuarzo, tiene poca resistencia a la tensión y lo destrozará un golpe fuerte. Por otro lado, el oro y el cobre pueden batirse en hojas; el cobre y la plata pueden estirarse en alambres.

COMPORTAMIENTO DE LOS MINERALES

El comportamiento de los minerales –las fuerzas internas que poseen y la forma en que reaccionan ante fuerzas o formas de energía externas– proporciona a menudo una pista para hallar su identidad. Entre las energías internas que distinguen un mineral de otro están el magnetismo y la radiactividad. Determinados minerales, como los compuestos de hierro magnetita y pirrotita, son magnéticos. Esta propiedad fue conocida por los antiguos. Las cualidades magnéticas de la magnetita o piedra imán fueron mencionadas por el filósofo griego Tales, que vivió en los siglos VII y VI a.C. La radiactividad, propia de los minerales uranio, torio y radio, no se descubrió hasta los últimos años del siglo XIX.

La luz es afectada en varias formas a medida que pasa a través de diferentes cristales. Los minerales trasparentes o traslúcidos pueden cambiar la velocidad y dirección de los rayos de luz, desviándolos, o refractándolos, a medida que éstos los atraviesan. Un cristal, como el de turmalina, puede polarizar un rayo de luz, o sea, forzarlo a vibrar en una dirección única o un único plano. Ciertos cristales dividen un rayo luminoso en dos, cada uno de los cuales vibra en un plano diferente. El mineral calcita tiene este efecto notable sobre la luz; produce

Pirita, u "oro de los tontos", es un sulfuro de hierro. Tiene cristales cúbicos macizos, de color amarillo-bronce y brillo metálico. A menudo se lo confunde con el oro.

C. Nuridsany

Katherine H. Jensen

Legrandita, un mineral de arsénico y cinc, exhibe la disposición radial de sus cristales alargados.

una imagen doble de cualquier cosa vista a través de él.

En ciertas circunstancias, algunos cristales emitirán luz. La piedra de Bolonia, una variedad del mineral baritina, o sulfato de bario, brillará en la oscuridad después de haber sido calentada ligeramente. Esta propiedad se denomina *termoluminiscencia*, o "luz por la aplicación de calor". La exhiben otros varios minerales, incluyendo la fluorita.

Algunos minerales dan destellos cuando se los golpea o frota suavemente. Esto se denomina *triboluminiscencia*, o "luz por fricción". La espalerita, o sulfuro de cinc, y el corindón de Carolina del Norte, u óxido de aluminio, responden de esta forma cuando se pasa una punta metálica por sus superficies.

Los destellos de luz emitidos en la triboluminiscencia no deben confundirse con las chispas que se producen al golpear varios metales, o el pedernal, una variedad del cuarzo.

Cuando determinados minerales se exponen en la oscuridad a rayos ultravioleta, brillarán en respuesta. Si no hay brillo al cesar la acción de los rayos, el fenómeno se denomina *fluorescencia*. Este nombre deriva del mineral fluorita, que exhibe acentuadamente esta propiedad. En ciertos casos, el brillo continuará un tiempo aun después que los rayos ultravioleta dejen de actuar contra la superficie del mineral. En dichos casos decimos que el mineral muestra la propiedad de *fosforescencia* (por el fósforo, que brilla en la oscuridad continuamente). La fluorescencia y la fosforescencia pueden estimularse por medio de rayos X y fuerte luz solar, como también por luz ultravioleta.

La emisión de luz de los tipos recién mencionados se ve afectada por las impurezas de un mineral dado. De acuerdo con éstas, el mineral calcita puede ser triboluminiscente, fluorescente o fosforescente. La luz que emiten diferentes minerales puede también mostrar variaciones de color.

Es posible hacer que ciertos minerales conduzcan electricidad o atraigan otros cuerpos eléctricamente. Cuando se calienta y luego se enfría, un cristal alargado de turmalina se cargará eléctricamente. Tendrá una carga positiva en un extremo del cristal y una carga negativa en el otro. Esto se denomina *piroelectricidad*, o "electricidad por calentamiento". Puede ilustrarse gráficamente por el siguiente experimento: un cristal de turmalina se carga de la manera descrita arriba. Luego se deposita sobre el cristal con un tamiz de seda, una mezcla pulverizada de plomo rojo y azufre. Las partículas adquieren cargas eléctricas opuestas a medida que pasan a través de la seda: las de plomo serán atraídas por un extremo del cristal; las de azufre, por el contrario.

La electricidad se produce en algunos cristales cuando se someten a presión mecánica. El cuarzo y la turmalina exhiben este comportamiento, denominado *piezoelectricidad* o "electricidad producida por presión". El efecto denominado *electroestricción*, o "presión por electricidad", es lo opuesto de la piezoelectricidad. En este caso, cuando una corriente eléctrica alternada pasa a través del cristal, éste responderá vibrando a tono con la corriente. El cristal tiene una frecuencia natural de vibración. Si las alternancias de la corriente son de esa frecuencia, las vibraciones del cristal se vuelven mucho más fuertes.

COLECCIONEMOS MINERALES Y ROCAS

Una colección de minerales es una de las más valiosas muestras de la naturaleza que podemos trasladar a nuestra casa. Los minerales son elementos o compuestos químicos que se producen de manera natural (dos o más elementos ligados químicamente). La mayoría de las rocas consiste en una mezcla de minerales.

DÓNDE ENCONTRARLOS

Para ampliar una colección de minerales es conveniente dirigirse al lugar donde la roca recién partida está al descubierto. Se pueden hallar excelentes ejemplares en una mina. Pero en ese lugar generalmente se despliega gran actividad, hay muchos trabajadores y máquinas y además puede ser peligroso. Nunca se debe penetrar en una mina sin el permiso de los propietarios y luego habrá que andar con mucho cuidado.

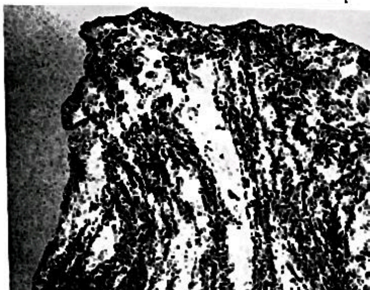
Hay otros lugares donde podemos hallar atractivos especímenes para nuestra colección. Una cantera de piedra es casi como una mina a cielo abierto, y en consecuencia debería brindar muchos buenos ejemplares. También vale la pena examinar los cortes que se han hecho para abrir caminos a través de las rocas. En esos lugares se hallarán vetas y filones donde los minerales han tenido oportunidad de crecer en grandes masas o aun en bolsones. Es muy posible que éstos estén recubiertos de cristales, que constituyen los especímenes de minerales más valiosos. Busquemos, pues, lugares en la roca donde haya concentraciones especialmente promisorias.

SEPARACIÓN DEL MINERAL

Una vez hallados los minerales habrá que separarlos de la roca. La mejor herramienta para este propósito es un pico de geólogo, un martillo con un extremo puntiagudo, hecho especialmente para exploradores y para coleccionistas de minerales como nosotros. Un cortafrío también es útil. Se recomienda usar gafas de seguridad para trabajar con martillo o cortafrío, a fin de proteger los ojos de las esquirlas que salten de la piedra. Si viajamos en automóvil, no sería una mala idea llevar una maza o martillo de clivaje destinado a las rocas grandes. Los bolsones con mejores revestimientos cristalinos se hallan casi siempre en el centro del mayor de los peñascos a la vista. Una vez de regreso a casa, podemos desbastar con un martillo más pequeño que el pico de geólogo las muestras recogidas hasta lograr el tamaño más conveniente.

Arriba: cuarzo humoso. Abajo: gneis.

C. Nuridsany



Llevemos una bolsa de recolección para colocar los especímenes que deseemos conservar, y algunos periódicos para envolverlos individualmente. No permitamos que ninguna pieza choque con otra, pues ambas podrían deteriorarse. Tampoco envolvamos las muestras demasiado pronto, hagámoslo al final, tomando sólo las mejores. Si vamos a más de una localidad en un día, o si no desenvolvemos las muestras al llegar a casa, anotemos en la parte exterior del papel de dónde proviene cada pieza, de modo que luego podamos roturarlas adecuadamente. No tomemos piezas demasiado grandes. Hay un límite para el número de muestras voluminosas que pueden incluirse en una colección.

IDENTIFICACIÓN DE MINERALES

El próximo paso es identificar las muestras. A tal fin, es conveniente usar una guía de campo de minerales. Hay alrededor de dos mil minerales conocidos; por tanto, sería casi imposible para los aficionados identificarlos. Debemos limitarnos a los minerales más comunes: unos 250 aproximadamente. Es posible identificarlos comparando sus propiedades con las enumeradas en una guía de campo o libro de texto. Damos aquí ciertas pautas para determinar las propiedades de las muestras.

Probamos la dureza de un mineral tomando trozos de minerales de prueba de dureza conocida y los usamos para ver si nuestro espécimen es más duro o más blando que aquéllos. Estos minerales de prueba forman una serie que se conoce como escala de dureza, o escala de Mohs. Al talco, que es el más blando de los conocidos, le corresponde el 1 en la escala. Al diamante, que es el más duro, el 10. A los demás minerales se les asignan números entre 1 y 10. Véase en esta obra el artículo "Minerales".

Para aplicar la escala, tratamos de rayar la muestra con uno de los minerales de prueba. Si no aparece ninguna rayadura en ella, esto significa que nuestro mineral es más duro que el de prueba. Se debe comenzar con el mineral de prueba más blando, pues si se comienza con el más duro, se podrá estropear el espécimen haciéndole varias rayaduras antes de determinar su dureza. Por otro lado, si se comienza con la piedra de talco y sucesivamente se toman los minerales que siguen en la escala, una simple rayadura dará la respuesta buscada.

Si una muestra desconocida es más dura que la fluorita (4 en la escala), pero puede rayarse con el mineral de prueba que lo sigue en la lista, decimos que su dureza es 4½. Si parece más cercana a la apatita (5 en la escala) que a la fluorita, podríamos atribuirle dureza de 4¾. Una vez que se consiguió rayar la muestra tratemos de borrar la marca con el dedo y observemos de nuevo. La muestra puede no haber sido rayada en absoluto, y la marca puede obedecer a polvo desprendido del mineral de prueba.

Si aún no estuviéramos seguros de la identidad de un mineral después de someterlo a todas las pruebas mencionadas, quizás haya que realizar un pequeño trabajo de laboratorio. El producto químico más empleado por los coleccionistas es ácido clorhídrico. Lo usamos, por ejemplo, para ayudarnos a determinar si una muestra es de calcita, mineral muy común. Calcita es carbonato de calcio ($CaCO_3$). El mármol, la piedra caliza común y la tiza son todas formas de esta sustancia. Si colocamos una gota de ácido clorhídrico sobre calcita se formarán burbujas a medida que parte de la calcita se disuelve y escapa el gas bióxido de carbono del mineral. Si el ácido está en un tubo de ensayo, se puede colocar un grano de calcita en el tubo y se formarán burbujas. Otros ácidos, tales como el acético, que se encuentra en el vinagre blanco, también hacen burbujear la calcita.

El mármol es una roca metamórfica. Es piedra caliza que ha cristalizado por deformación. Arriba: mármol blanco. Abajo: mármol negro. Derecha: una cantera.

Bevilacqua-Cedri

Bevilacqua-Cedri

En otro tipo de ensayo hacemos una versión en miniatura de un horno de fundición. La fundición consiste en calcinar minerales con el propósito de separar los metales deseados de los elementos químicos con que están combinados. Por ejemplo, elementos metálicos, tales como cobre, plomo o cinc, a menudo se combinan en estado natural con azufre, constituyendo los compuestos químicos denominados sulfuros. Muchos de ellos sólo deben calcinarse en la llama caliente de un fundidor para quemar el azufre, dejando el metal.

El recolector de mineral suele tener un horno de fundición privado, consistente en un soplete y un trozo de carbón vegetal en el que se ha practicado un pequeño hoyo. Coloquemos en el hoyo una pizca del mineral que consideremos una mena metálica. Luego usemos como fuente de calor una vela o llama de alcohol. Por medio del soplete, dirijamos la llama dentro del hoyo del carbón. Si la muestra es un sulfuro, saldrán vapores de azufre. Posiblemente se forme una pequeña bola de metal caliente al rojo, que girará en el hoyo del carbón como si fuera una perla viviente.

Remy-Atlas-Photo

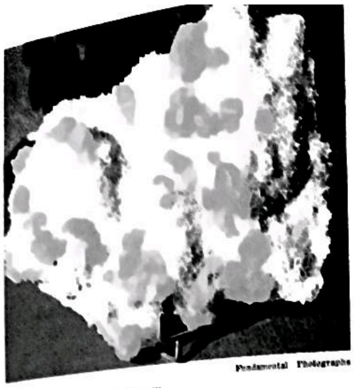
Cuando la bola metálica se enfría podemos achatarla con un martillo y ver si se parece al plomo, la plata o el cobre. Si logramos identificarla con alguno de éstos, conoceremos dos de los elementos de nuestro mineral: el azufre y el metal.

REGISTREMOS LAS MUESTRAS

Después de haber identificado nuestro espécimen debemos roturarlo y quizá darle entrada en un catálogo. El rótulo debe tener el nombre del mineral y del lugar donde se recogió. También sería útil agregar la fecha. Muchos aficionados numeran además la muestra y llevan un archivo de tarjetas en el que colocan información adicional de interés, tal como forma cristalina, o los otros minerales que conocen en la muestra. Una numeración mucho mejor es usar los números de especímenes empleados en un libro de texto de uso común y agregar los propios. Cualquiera sea el sistema que usemos, es menester llevar algún tipo de registro. Una acumulación de minerales sin ninguna clasificación no tiene casi valor para nadie que no sea el propio coleccionista, y aun para él su valor es limitado.

GEMAS

Entre las muchas variedades de minerales hallados en la Tierra, algunos son particularmente estimados porque son raros, o duraderos, o hermosos, o porque combinan estas cualidades. Dichos minerales se dividían antes en dos clases principales. Aquellos particularmente destacados, tales como diamantes, esmeraldas y rubíes, se conocían como *piedras preciosas*. Otros que no eran tan raros ni tan hermosos se denominaron *piedras semipreciosas*; incluyen ortosas, circón y ágata. Esta distinción aún se hace a veces. Generalmente, sin embargo, cualquier mineral que posea señaladas cualidades de belleza y durabilidad y que pueda usarse como adorno personal después de haberse labrado se denomina *gema* o piedra gema. Algunos coleccionistas de minerales practican el *hobby* de cortar y pulir distintas piedras gemas que encuentran en el campo. El arte de pulir



Fundamental Photographs

Roca sulfurosa de Sicilia.

gemas se llama lapidación y lo realizan personas a las que se denomina lapidarios.

ROCAS

Cuando hayamos adquirido un conocimiento aceptable de alguno de los minerales más importantes, quizá queramos comenzar una colección de rocas. Antes de hacerlo, deberíamos poseer algún conocimiento de los diferentes tipos de rocas que constituyen la corteza terrestre.

Hay tres tipos principales de éstas: ígneas, sedimentarias y metamórficas. Las ígneas se formaron por la solidificación de materia fundida productora de roca. De esta materia, alguna fluyó de los volcanes y se enfrió muy rápido. Los granos aislados del mineral no tuvieron muchas posibilidades de crecer. La roca ígnea que resultó es de grano fino y compacta. Si la roca fundida ha permanecido bajo la superficie terrestre hasta solidificarse, se enfrió lentamente. Los cristales contaron con mucho tiempo para formarse y crecieron mucho. El resultado es una roca ígnea de grano grueso, como el grafito. Tales rocas han quedado expuestas generalmente por la elevación de la corteza o la erosión de las rocas superpuestas.

Las rocas sedimentarias consisten en sedimentos trasformados en piedra por el trascurso de los siglos. Derivan no sólo de fragmentos de roca, sino también de restos de plantas y animales. En algunos casos se han formado por la evaporación del agua oceánica. Las capas sedimentarias de la corteza terrestre constituyen gran parte de su superficie. En algunos lugares, la acumulación de capas puede tener muchos cientos de metros de espesor.

Las rocas metamórficas son antiguas rocas ígneas o sedimentarias que han sido trasformadas dentro de la corteza terrestre en tipos totalmente diferentes. Esta trasformación se debió a distintos factores, tales como presión, temperaturas extremadamente altas y la presencia de agua. Las partículas originales de roca han sido forzadas a adoptar nuevas disposiciones, dando lugar algunas veces a la formación de nuevos minerales.

He aquí un breve resumen de algunas formaciones de rocas importantes que podrían hallarse en el trascurso de una recolección.

Rocas ígneas: de grano parejo, normalmente sin capas o vetas pronunciadas, consistentes en minerales de silicato y, por tanto, relativamente duras y compactas. Entre las rocas ígneas de grano fino están: (1) Basalto: una roca oscura (negra, castaño, verde oscuro o gris oscuro) de origen volcánico. (2) Pórfidos: rocas con grandes cristales engastados en una masa de otros de grano más fino, denominada pasta. (3) Riolitos: rocas de colores comparativamente claros, que van del blanco al gris, al rosado y al púrpura.

Entre las rocas ígneas de grano grueso están: (1) Granitos: compuestos principalmente de cuarzo y feldespato y, generalmente, también de mica. (2) Dioritas: compuestos de feldespato y uno o más minerales oscuros; el feldespato predomina. (3) Gabros: compuestos de feldespato y uno o más minerales oscuros; predominan estos últimos. (4) Peridotitas: generalmente negras u oscuras; contienen grandes cantidades de hierro. (5) Pegmatitas: granitos de grano particularmente grueso, con grandes cristales de cuarzo, feldespato y mica.

Rocas sedimentarias: constituidas por fragmentos minerales más o menos redondos. Las rocas sedimentarias generalmente muestran estratificación, si no en el espécimen recolectado, seguramente en la cantera o en el corte donde éste fue hallado.

Hay cinco tipos de rocas sedimentarias: (1) Conglomerado: guijarros o fragmentos de roca más grandes cementados entre sí. (2) Piedra arenisca: granos de arena cementados entre sí. (3) Pizarra: arcilla que se ha convertido en masas de roca en láminas. (4) Piedra caliza: compuesta de conchillas microscópicas, o carbonato de calcio precipitado, o ambos. Muchas

piedras calizas tienen arcilla y arenas mezcladas con ellas. (5) Yeso, sal común, sal de Glauber y sales de Epsom: compuestos contenidos en el agua de mar que precipitaron a medida que ésta se evaporó.

Rocas metamórficas: contienen algunos de los mismos minerales que el granito, pero en distintas proporciones. Las rocas metamórficas pueden consistir en una sola clase de mineral y no en una mezcla de varios, como en muchas rocas ígneas.

Las rocas metamórficas son: (1) Gneis: se parece mucho al granito, pero tiene capas alternadas de diferentes minerales que le dan apariencia veteada; la mayoría de las variedades contiene mica negra. (2) Esquistos: muy parecidos al gneis, pero el mineral más común puede ser mica; puede ser dividido en láminas. (3) Cuarcita: no parece muy distinta de la piedra arenisca de la cual proviene, pero está más fuertemente cementada. Cuando un trozo se quiebra, los granos de arena se parten al través, en lugar de alrededor de los mismos, como en el caso de la piedra arenisca. La cuarcita es una de las rocas más duras y más resistentes. Un afloramiento (superficie expuesta) de un lecho de cuarcita es probable que tome la forma de un risco. (4) Mármol: constituido por la recristalización de la piedra caliza; puede ser fino o bastante grueso, y de casi cualquier color. Puesto que es calcita, puede rasparse con un cuchillo y burbujea por la acción del ácido. (5) Pizarra: de grano tan fino que los granos individuales del mineral no pueden distinguirse fácilmente. Se parte en planchas lisas, la mayoría de color negro azulado, pero también hay variedades negras, verdes y rojas.

FÓSILES

Los recolectores de rocas pueden hallar a veces fósiles en formaciones sedimentarias, particularmente en piedras calizas y pizarras. Hay varios tipos de fósiles. Los más comunes son los caparazones de distintos animales. Es más probable hallarlos en los lechos de piedra caliza ligeramente esquistosos, especialmente en las pendientes desgastadas por la intemperie. Los fósiles son a menudo caparazones reemplazados por calcita o incluso por cuarzo. Los caparazones fósiles pueden sobresalir en una pendiente o estar encaramados en un pedestal, como pelotas de golf listas para ser lanzadas.

Las hojas fósiles se hallan como impresiones entre dos capas de pizarra. Hay que partir y separar las lajas para hallar los fósiles. Uno de los mejores lugares para buscarlas es en los esquistos superpuestos a lechos de carbón.

Los animales terrestres también han dejado restos fósiles. Como dichos animales no se desarrollaron hasta tarde en la historia de la Tierra, el número y variedad de sus fósiles no puede compararse con los de animales marinos.

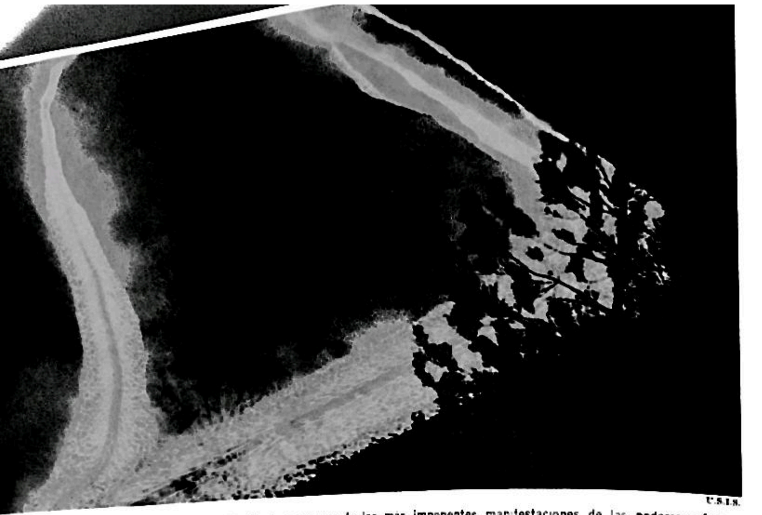
GUÍAS PARA COLECCIONISTAS

Cuando hayamos llegado a ser coleccionistas expertos de rocas y minerales nos resultará provechoso suscribirnos a revistas que traten de nuestro hobby, como también informarnos sobre comerciantes de minerales y clubes de coleccionistas.

Coleccionemos lo que podamos, pero no vacilemos en llenar cualquier brecha de la colección recurriendo a comerciantes. Esto es más barato que internarse en lugares intrincados a fin de buscar muestras faltantes.

Las lavas son ejemplos de rocas ígneas. De izquierda a derecha: basalto, obsidiana, escoria.

Izquierda y derecha: A. W. Ambler, National Audubon FR; centro: [illegible], National Audubon FR



U.S.I.S.

El vulcanismo es una de las más imponentes manifestaciones de las poderosas fuerzas que actúan en lo profundo del interior de la Tierra.

EL INTERIOR DE LA TIERRA

El hombre ha llegado a conocer muchas cosas sobre la Tierra mediante la observación directa o, al menos, bastante directa. La ha explorado, inspeccionado y cartografiado desde el aire, y analizado sus rocas. No podríamos aplicar estos métodos al estudio de su interior. Las minas más profundas penetran en éste menos de tres kilómetros, y las perforaciones no van mucho más allá; son meros alfilerazos en un planeta de la dimensión del nuestro.

Por tanto, se deben usar medios indirectos para conocer la estructura interna de la Tierra. El geofísico reúne datos de varias fuentes, analiza los que aportan los terremotos y la rotación terrestre, mide las mareas y estudia las variaciones de la gravedad en distintos lugares de la superficie. También trata de reproducir en el laboratorio las condiciones que se supone existen en el interior de nuestro planeta.

A causa de la naturaleza indirecta de las pruebas con que trabaja, el geofísico es cauteloso en la mayoría de sus asertos y prefiere usar la palabra "probablemente" en muchas oportunidades. No obstante, no debemos olvidar que hay gran cantidad de razonamiento probable en toda clase de investigación científica. Todos los juicios humanos son inseguros, al menos hasta cierto punto, aun cuando se basen en testimonios directos. A decir verdad, los instrumentos científicos que usan los geofísicos son, frecuentemente, más de fiar que el ojo humano.

Algunas conclusiones pueden aceptarse con certeza. Es tan probable, por ejemplo, que exista un núcleo central denso en la Tierra, como que el Sol aparezca mañana. En otros casos, los descubrimientos de los geofísicos son tentativos, y las teorías deberán modificarse según la nueva información que se adquiera.

LOS TERREMOTOS PROPORCIONAN INDICIOS

La misma Tierra provee los medios principales para resolver el misterio de su estructura interna. Todos los años ocurre una cantidad de terremotos en muchas partes del globo. Cada uno libera súbitamente una cantidad enorme de energía, la que viaja en forma de ondas desde el *foco*, o fuente de origen de la perturbación, a través de todas las partes de la Tierra, incluyendo las más profundas. Cuando las ondas emergen a la superficie las registran instrumentos llamados *sismógrafos*, instalados en varios miles de observatorios.

La aguja del sismógrafo traza los movimientos de la Tierra bajo nuestros pies, en registros llamados *sismogramas*. Si un terremoto ocurre, digamos, en Nueva Zelanda, alrededor de 20 minutos después un observador en Inglaterra puede ver cómo la aguja del sismógrafo traza el diagrama de las ondas que están llegando.

Algunos sismólogos del siglo XX han preparado tablas exactas, que dan el tiempo de desplazamiento de las ondas sísmicas a través de todas las zonas de la Tierra. Las tablas de Jeffreys-Bullen ahora se usan internacionalmente. Fueron preparadas por sir Harold Jeffreys, de la Universidad de Cambridge, con la ayuda de Keith E. Bullen, durante el período 1931-1940. B. Gutenberg y C. F. Richter, del Instituto de Tecnología de California, recopilaron otras tablas muy útiles.

Dos tipos de ondas penetran profundamente en la Tierra: las ondas P (ondas primarias) y las ondas S (ondas secundarias). Las ondas P viajan a través de las partes sólidas y fluidas del interior de la Tierra. En las rocas próximas a la superficie avanzan a unos 5 kilómetros por segundo, alcanzando su máxima velocidad de 14 kilómetros por segundo a una profundidad de unos 2.800 kilómetros. Las ondas S se desplazan a unos dos tercios de la velocidad de las ondas P en regiones sólidas, y no pueden pasar a través de zonas líquidas o fluidas. Cuando las ondas llegan a un límite entre dos estratos pueden ser reflejadas, dirigiéndose hacia arriba, en dirección a la superficie, o bien ser refractadas o curvadas a medida que pasan a un estrato próximo. Esto significa que penetran en este estrato cambiando su dirección. Las ondas que alcanzan la corteza terrestre son nuevamente reflejadas hacia abajo.

La velocidad de las ondas P y S varía según la profundidad de las zonas por las que se desplazan. El sismólogo puede calcular todas estas variaciones y trazar luego una gráfica, dividiendo el interior de la Tierra de acuerdo con su profundidad.

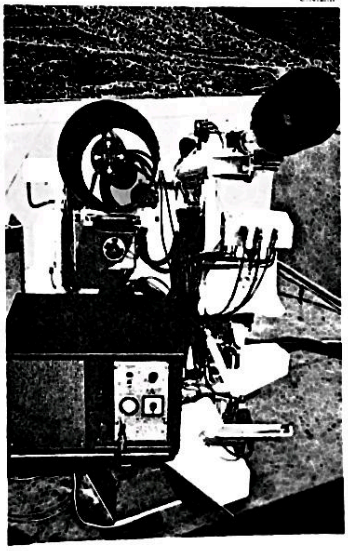
Los sismogramas pueden compararse a los rayos X en el cuerpo humano. La estructura interna del cuerpo influye sobre la intensidad de estos rayos a medida que pasan entre su fuente y la película o placa fotográfica. En cuanto a las regiones más profundas de la Tierra, afectan las ondas sísmicas que viajan entre el punto de origen de un terremoto y el sismógrafo. Es mucho más difícil interpretar los sismogramas que las fotografías de rayos X. Si un médico examina una radiografía ve una imagen que presenta una exacta similitud con el cuerpo humano. Un sismograma, en cambio, muestra solamente un esquema intrincado de líneas onduladas. Para descifrarlas, el geofísico necesita recurrir a los métodos de las matemáticas y la física.

DOS REGIONES INTERIORES

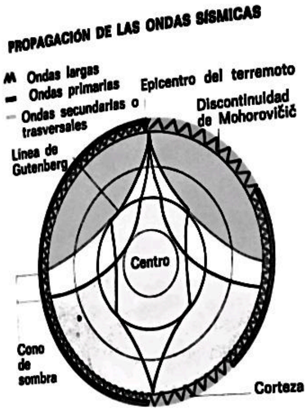
Los datos obtenidos de los sismogramas indican que el interior de la Tierra está dividido

Los geólogos han perfeccionado muchos instrumentos delicados para estudiar la Tierra. Partiendo de los datos recogidos por estos aparatos, se pueden reunir ideas concernientes al interior terrestre. Esta foto muestra un aparato usado para estudiar el magnetismo de la Tierra.

C.N.E.S



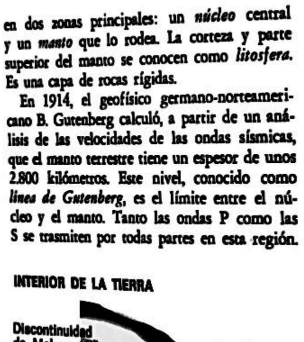
PROPAGACIÓN DE LAS ONDAS SÍSMICAS



en dos zonas principales: un *núcleo* central y un *manto* que lo rodea. La corteza y parte superior del manto se conocen como *litosfera.* Es una capa de rocas rígidas.

En 1914, el geofísico germano-norteamericano B. Gutenberg calculó, a partir de un análisis de las velocidades de las ondas sísmicas, que el manto terrestre tiene un espesor de unos 2.800 kilómetros. Este nivel, conocido como *línea de Gutenberg*, es el límite entre el núcleo y el manto. Tanto las ondas P como las S se trasmiten por todas partes en esta región.

INTERIOR DE LA TIERRA



Puesto que las ondas S no atraviesan la materia fluida, el manto debe ser esencialmente sólido en su conjunto, excepto en las zonas oceánicas y en limitados depósitos de materia volcánica.

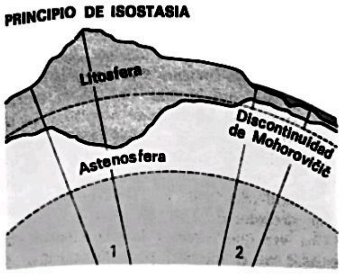
En 1909, A. Mohorovicic hizo un significativo descubrimiento al analizar los sismogramas de un terremoto en los Balcanes. Encontró que había un notable cambio en las velocidades de las ondas sísmicas después de haber penetrado unas decenas de kilómetros bajo la superficie de la Tierra. Las ondas P y S viajaban a velocidades más lentas y más variables antes de alcanzar este nivel que posteriormente. El nivel corresponde a un límite más o menos definido conocido como *discontinuidad de Mohorovicic.*

Tiempo después se descubrió que este límite se extiende por todo el mundo y que, generalmente, se halla a 35 kilómetros bajo la superficie terrestre en las zonas continentales. La profundidad es algo mayor bajo las grandes cordilleras; y es sólo de unos 8 kilómetros bajo el fondo de los océanos más importantes. La parte del manto situada por debajo de la litosfera y encima de la discontinuidad de Mohorovicic se llama *atenosfera.* Es más blanda y más fácilmente deformable que la capa superior. El manto se vuelve otra vez rígido debajo de la atenosfera.

El núcleo de la Tierra se divide en tres regiones: una interior, una exterior y un estrato de transición de 480 kilómetros entre ambas. Las ondas S no pueden penetrar la materia líquida, y el hecho de que no atraviesen el núcleo externo hace creer que éste se encuentra en estado líquido, conclusión bien establecida por otras pruebas. Las ondas P entran en el núcleo. Al hacerlo, su velocidad desciende súbitamente de 13,5 kilómetros por segundo a 8 kilómetros por segundo, y son bruscamente refractadas. Esta refracción reduce mucho la intensidad de las ondas que penetran en la región exterior del núcleo, y atraviesan una capa de su superficie llamada zona de sombra. Fue la presencia de esta zona lo que primero proporcionó a los geofísicos la prueba de que realmente existe un núcleo central. Hay un marcado aumento de velocidad en las ondas P cuando alcanzan el núcleo interno, importante indicio de que esta parte del núcleo es sólida.

El núcleo interno tiene un radio de unos

PRINCIPIO DE ISOSTASIA



Todas las regiones de la corteza terrestre están en un estado de equilibrio conocido como *isostasia*. Si las capas exteriores de la Tierra se dividieran en dos columnas (una en una zona montañosa y la otra bajo el mar), estas columnas tendrían el mismo peso. Las diferencias en el peso de la litosfera serían compensadas por la deformable atenosfera.

1.300 kilómetros. El espesor del núcleo externo supera los 1.600 kilómetros. Por tanto, el radio total de los núcleos interior y exterior combinados, más el estrato de transición, es de unos 3.500 kilómetros o algo menos.

El interior de la Tierra ha sido dividido, tentativamente, en varias otras regiones. En especial, hay signos de algunos cambios bastante marcados en las rocas que están bajo la corteza terrestre.

DENSIDADES Y PRESIONES

Podemos hallar la densidad de cualquier cuerpo si conocemos su masa y su volumen, porque la densidad es igual a la masa por unidad de volumen. Se sabe que la masa de la Tierra es aproximadamente de 6×10^{21} toneladas. Su volumen es más o menos de $1{,}1 \times 10^{12}$ kilómetros cúbicos. Mediante estos datos, los científicos han calculado que la Tierra tiene una densidad promedio de $5{,}517 \pm 0{,}001$ gramos por centímetro cúbico. El agua tiene una densidad de 1 gramo por centímetro cúbico a la temperatura ordinaria. Esto significa que la Tierra es aproximadamente 5,517 veces más densa que el agua. La densidad de una sustancia comparada con la del agua se denomina *densidad relativa.* La cifra 5,517 representa, por tanto, la densidad relativa media de la Tierra.

La mayoría de las rocas en la superficie terrestre tiene una densidad relativa menor de 3. Por supuesto, desde que la densidad relativa del total de la Tierra es más de 5, la parte interior de ella debe contener materiales más densos que los de las rocas superficiales. Por medio del estudio de las gráficas de ondas sísmicas, los geólogos pueden calcular las densidades de las distintas partes del manto y del núcleo, ya que la velocidad de las ondas en una región depende de la densidad de los materiales que van atravesando.

Otras clases de pruebas también son necesarias para calcular las densidades de varios puntos bajo la superficie terrestre. Una fuente de información está basada en la rotación del planeta. Sabemos que, en términos generales, la Tierra gira alrededor de un eje que pasa a través de los dos polos y que, en parte como resultado de esta rotación, se ha producido cierto abultamiento de nuestro planeta

En un intento de conocer cómo el interior de la Tierra está afectando a la corteza, los científicos han tomado "testigos" del suelo oceánico usando equipos como esta "super bombilla".

Institución Oceanográfica Woods Hole

Keystone

La Tierra como se la representaba en el siglo XV. Hemos aprendido mucho desde entonces, pero aún hay numerosos interrogantes, sobre todo los relacionados con la naturaleza de su núcleo.

en el ecuador. La atracción gravitatoria de la Luna sobre el abultamiento ecuatorial causa ciertos cambios pequeños en la dirección del eje, que los astrónomos han descubierto y medido. Los geofísicos estudian estos cambios en relación con ciertos hechos conocidos relativos a la forma de la Tierra. Varias teorías físicas de elasticidad y de atracción gravitatoria también ayudan a determinar las densidades, así como los experimentos sobre las rocas.

Usando todos estos medios, Keith E. Bullen calculó, en 1936, que la densidad relativa de la Tierra va desde un poco más de 3, exactamente debajo de la corteza, hasta alrededor de 5,5 en la base del manto. Este aumento de densidad con la profundidad es producido, principalmente, por la tremenda presión ejercida por las capas de rocas suprayacentes. También puede deberse, en parte, a cambios en la composición química del manto.

En el límite entre el manto y el núcleo, la densidad relativa salta bruscamente de 5,5 a cerca de 10. Dentro del núcleo exterior, la densidad relativa aumenta a causa de la mayor presión, hasta que alcanza 11,5 en el fondo. Estas cifras son ampliamente aceptadas en la actualidad.

La densidad del núcleo interior ha sido calculada exactamente. El Dr. Bullen estimó que la densidad relativa de la Tierra en su centro es de alrededor de 13.

Una vez que sabemos la variación de la densidad en el interior, podemos calcular la distribución de presiones. La presión atmosférica es de alrededor de un kilogramo por centímetro cuadrado al nivel del mar. En el lecho del océano Pacífico, la presión es algo menos de una tonelada por centímetro cuadrado. Esta impresionante cifra es insignificante si se la compara con la presión en el profundo interior de la Tierra. El Dr. Bullen ha calculado que la presión en el fondo del manto es de unas 1.550 toneladas por centímetro cuadrado y de unas 3.900 toneladas por centímetro cuadrado en el centro de la Tierra. Se necesitaría una columna de acero de unos 800 kilómetros de altura apoyada sobre un centímetro cuadrado de superficie para producir una fuerza de 3.900 toneladas.

ELASTICIDAD DE LA TIERRA

La Tierra no es un cuerpo absolutamente rígido, sino que reacciona elásticamente ante las tensiones, exactamente como el acero. Por esta razón, la atracción del Sol y de la Luna provoca movimientos de marea, no sólo en los océanos sino también en tierra firme. Ciertas fluctuaciones en la posición del eje terrestre se deben también a la elasticidad de nuestro planeta. La *rigidez* de un cuerpo es su resistencia a las fuerzas que tienden a deformarlo. Lord Kelvin, en 1863, calculó, sobre la base de estos efectos, que el promedio de rigidez de la Tierra era bastante mayor que el del acero. Esto significa que, considerado en conjunto, nuestro planeta tiene una rigidez mucho mayor que la del promedio de las rocas de superficie. Desde la época de Kelvin, el conocimiento de las variaciones de la densidad terrestre hizo posible calcular la rigidez y las velocidades sísmicas del interior de la Tierra. La rigidez aumenta a través de todo el manto, y el Dr. K. E. Bullen estimó que en el fondo de éste esa rigidez es unas tres veces mayor que la del acero común.

El científico japonés H. Takeuchi usó los datos sobre terremotos y movimientos de mareas de la tierra firme para demostrar que la mayor parte del núcleo exterior es mucho menos rígida que el manto. Este descubrimiento da el más fuerte apoyo a la teoría de que el núcleo exterior está en estado de fusión. Una serie independiente de cálculos del Dr. Bullen han probado que es altamente probable que el núcleo interior sea sólido. Esta parte del núcleo central tiene una rigidez comparable a la del acero común.

TEMPERATURA Y FLUJO DE CALOR

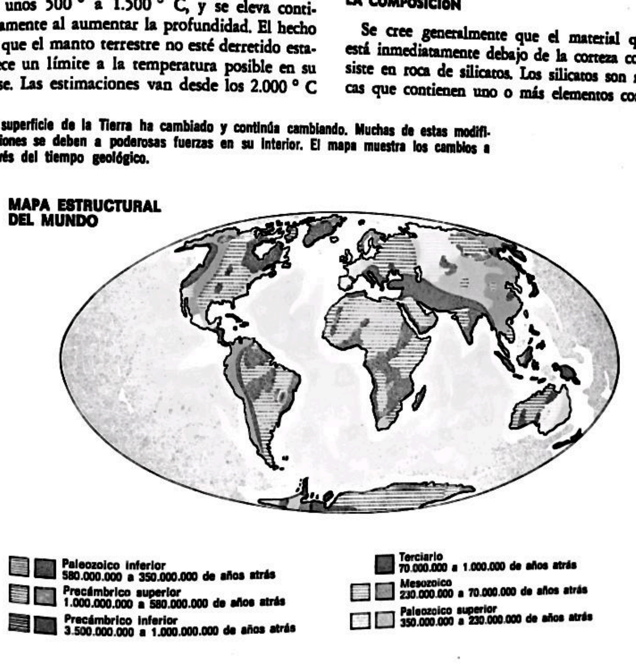
No tenemos datos precisos acerca de la temperatura en el interior de la Tierra. La temperatura en el fondo de la corteza es, probablemente, de unos 500 ° a 1.500 ° C, y se eleva continuamente al aumentar la profundidad. El hecho de que el manto terrestre no esté derretido establece un límite a la temperatura posible en su base. Las estimaciones van desde los 2.000 ° C hasta cerca de 6.000 ° C. La temperatura en el centro de la Tierra posiblemente supere en menos de 1.000 ° C estas cifras.

Hay un pequeño pero constante desprendimiento de calor proveniente de la Tierra; la mayor parte de este calor está causado por los materiales radiactivos de rocas próximas a la superficie. En las regiones continentales, la radiactividad se halla, principalmente, en rocas graníticas de la corteza, las que no se encuentran en la mayoría de las regiones oceánicas. Se suponía, por tanto, que el calor despedido por los fondos oceánicos podría ser mucho menor que el de los continentes. Las últimas mediciones, sin embargo, parecen indicar que el flujo de calor desde la Tierra es bastante uniforme sobre toda su superficie. Las fluctuaciones no exceden el 20 por ciento.

LA COMPOSICIÓN

Se cree generalmente que el material que está inmediatamente debajo de la corteza consiste en roca de silicatos. Los silicatos son rocas que contienen uno o más elementos com-

La superficie de la Tierra ha cambiado y continúa cambiando. Muchas de estas modificaciones se deben a poderosas fuerzas en su interior. El mapa muestra los cambios a través del tiempo geológico.



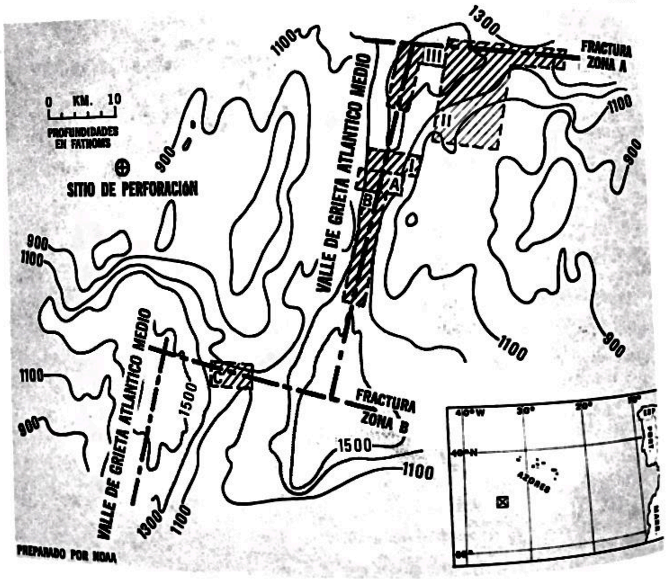
binados con silicio y oxígeno. No hay certeza sobre el tipo de roca que predomina en esta región; sin embargo, se cree probable que el mineral olivina, que es un silicato de hierro y magnesio, se halle ahí en cantidad. Esto se prueba comparando las velocidades de las ondas sísmicas para esta región con ciertos experimentos hechos sobre rocas en laboratorios de geofísica.

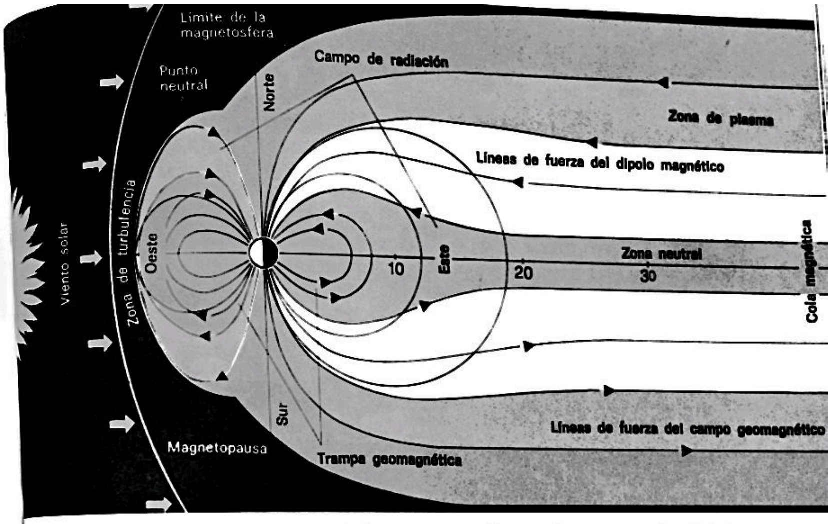
En la parte inferior del manto, la composición de las rocas parece ser casi la misma que más arriba. Es posible, sin embargo, que en esta parte del manto el material consista en fases bien discernibles de sílice, magnesio y óxido de hierro, según lo ha sugerido Francis Birch, de la Universidad de Harvard.

Hasta hace muy poco tiempo, era creencia casi universal que el núcleo central de la Tierra estaba formado principalmente de ferroníquel. Esta idea se apoyaba, en buena medida, en el estudio de los meteoritos –fragmentos caídos sobre la superficie terrestre desde el espacio exterior– que pueden ser de tres clases. Están los "sideritos", que consisten en hierro frecuentemente mezclado con níquel; los "lititos", que se asemejan a rocas de la superficie terrestre, y los "sideritos pétreos", una mezcla de los dos tipos anteriores.

Hoy, generalmente, se está de acuerdo en que los meteoritos son fragmentos de cuerpos mucho mayores. Un análisis de los sideritos indica que los cristales que contienen se formaron a través del lento enfriamiento de ferroníquel sólido, lo cual ocurrió a presiones existentes solamente en la profundidad de cuerpos de considerable tamaño. Se ha calculado que los sideritos se formaron dentro de un cuerpo quizá tan grande como la Luna.

La Cadena Mesoatlántica ha sido el asiento de algunos de los estudios más interesantes sobre lo que está sucediendo ahora en la Tierra. Se la cree un lugar donde la aportación del interior de ésta a la formación de la corteza puede ser observada y medida.
NOAA





La Tierra está rodeada por una vasta envoltura conocida como magnetosfera. El viento solar –la constante corriente de partículas emitidas por el Sol– modela esta envoltura como muestra la ilustración superior.

Esto sugiere que, por lo menos, algunos meteoritos han venido de un planeta que fue alguna vez como la Tierra, pero que después estalló en fragmentos. Los fragmentos de semejante cuerpo serían como los materiales de las profundidades de la Tierra. Si, como se creía, la mayoría de los meteoritos fueran sideritos, el ferroníquel debería predominar en el núcleo del planeta.

Hoy día sabemos que la proporción de sideritos con respecto a los lititos no es tan grande como se suponía. Una reciente investigación de J. Öpik bajó la proporción a un dos por ciento en masa. Otros investigadores estimaron porcentajes de sideritos más altos que Öpik. Sin embargo, la cuestión está lejos de haber sido resuelta. Si aceptamos la cifra de Öpik, o una poco mayor que la de él, resulta evidente que la prueba suministrada por los meteoritos para el contenido de ferroníquel del núcleo central no es tan concluyente como parecía antes.

La mayoría de los geofísicos cree todavía que el núcleo interior de la Tierra está formado por ferroníquel, con el agregado, quizá, de algunos materiales ligeramente más densos. Muchos sostienen, además, que el núcleo externo también consiste principalmente en ferroníquel, no sólido, sino derretido. Otros dudan. En 1948, el científico británico W. H. Ramsey sugirió la idea de que el núcleo exterior tenía la misma composición química que la parte inferior del manto. La diferencia de densidad entre las dos regiones, sostenía, se debe al aumento de las presiones bajo el manto. Esta teoría necesita muchas más pruebas antes que pueda ser considerada en reemplazo de la teoría anterior a la que nos referimos arriba. Sin embargo, ha influido sobre el pensamiento de algunos geofísicos.

EL CAMPO MAGNÉTICO DE LA TIERRA

En el año 1600, sir William Gilbet, un médico inglés, propuso la teoría de que la Tierra tiene las propiedades de un enorme imán, cuyos polos magnéticos casi coinciden con los polos geográficos. También sugirió que el campo magnético terrestre se origina, principal-

mente, en las profundidades del interior del planeta. Estas ideas han sido confirmadas por muchos investigadores.

Se han propuesto muchas teorías para explicar el magnetismo de nuestro planeta. Una suponía que había hierro permanentemente magnetizado en las profundidades de la Tierra. Esta teoría tuvo que descartarse cuando se demostró que el núcleo terrestre era parcialmente fluido y por tanto no podía retener un magnetismo permanente. Parece probable que el magnetismo permanente en cualquier otro lugar del interior de la Tierra no proporcionaría un campo magnético lo suficientemente fuerte.

En 1947, Patrick M. S. Blackett, de la Universidad de Londres, sugirió que cualquier cuerpo sólido en rotación, tal como nuestro planeta, genera un campo magnético simplemente porque gira. Pero los experimentos de laboratorio con objetos grandes en rotación no revelaron la existencia de ningún campo semejante. Una prueba sobre la Tierra misma, ideada por sir Edward C. Bullard y realizada por Keith Runcorn en minas profundas, fracasó también como sostén para la teoría de Blackett.

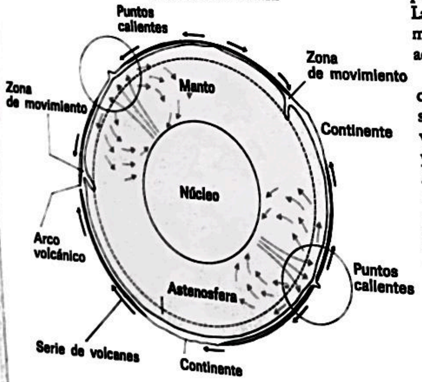
Ahora parece prácticamente seguro que el campo magnético de nuestro planeta está generado por corrientes eléctricas ordinarias que circulan por el interior de la Tierra. En 1893, sir Harold Lamb señaló que tales corrientes tendrían que ser continuamente alimentadas por alguna fuente de energía en el interior. Es natural suponer que esto debería suceder en la parte de la Tierra donde hay menor resistencia eléctrica, es decir, en la región exterior fluida del núcleo.

Walter M. Elsasser, un físico norteamericano nacido en Alemania, sugirió en 1939 que la corriente podría producirse en el núcleo cuando materiales de distintas propiedades eléctricas y temperaturas ligeramente diferentes se ponen en contacto. A ésta se la llama *hipótesis termoeléctrica*. La *termoelectricidad* se produce por el calentamiento desigual de un circuito eléctrico compuesto por dos metales distintos. En 1954, Runcorn sugirió que podría haber un efecto termoeléctrico en el límite entre el manto y la parte exterior del núcleo.

Actualmente, la teoría más importante sobre el magnetismo de la Tierra es la *teoría de la dínamo*, de Elsasser y Bullard. Ella sostiene que una enorme dínamo natural, en lo profundo de la Tierra, convierte energía mecánica en energía magnética; la energía mecánica sería suministrada por un tipo especial de movimiento del fluido, llamado *convección*, que trasporta corrientes eléctricas dentro del núcleo exterior. Elsasser hizo elaborados cálculos matemáticos para demostrar que ese movimiento es posible. La teoría de la dínamo sobre el origen del magnetismo terrestre está considerada en la actualidad la más promisoria de todas.

Poco a poco, está surgiendo una idea más clara del interior de la Tierra. Algo sabemos sobre las diferentes regiones en que está dividida y, también, sobre su densidad, presión y elasticidad a diversas profundidades. Tenemos alguna idea de la composición química de los materiales en las profundidades de nuestro planeta. Por lo menos, hemos dado un primer paso promisorio hacia la solución del problema de cómo se genera el campo magnético de la Tierra. Todavía queda mucho por aprender acerca de ella antes de que llegue el momento en que sepamos tanto de su interior como sabemos de su exterior.

Según teorías recientes, material del manto y del núcleo de la Tierra produce "focos calientes" y zonas especiales de movimiento cerca de la corteza.



Gianni Tortoli, P.R.

VOLCANES Y GÉISERES

Visión infernal del monte Etna, durante una erupción en 1971. La lava mana de tres fisuras y converge a medida que fluye por la ladera del volcán.

Los volcanes y los géiseres deben estudiarse juntos, pues están relacionados; sin embargo, los primeros son mucho más espectaculares que los segundos; en verdad, representan la manifestación más imponente de las fuerzas naturales, con la posible excepción de los terremotos. En todo el mundo, los cráteres volcánicos arden, humean y rugen. Ocasionalmente, desvastan la tierra con nubes de fuego y vapor, con andanadas de peñas y rocas, y con ríos de lava y fango caliente.

Un volcán es una chimenea, o fisura, en la corteza terrestre, desde donde, a intervalos, emergen gases, líquidos y sólidos calientes con mayor o menor violencia. Los materiales que surgen de la chimenea forman, tarde o temprano, un *cono* en torno de ella. Con frecuencia, el término volcán no sólo se aplica a la chimenea, sino a la montaña misma. Las montañas volcánicas difieren grandemente en altura y forma. En general son cónicas y, a menudo, muy elevadas; Orizaba y Popocatepetl, en México; Cotopaxi, en los Andes; Monte San Elías, en Alaska; Kilimanjaro y Kenya, en África; y Ararat y Demavend, en Asia, tienen todas entre 5.000 y 7.000 metros de altura.

Muchos de los picos cónicos de los volcanes forman rasgos característicos inolvidables; entre ellos el Pico de Tenerife, que surge del mar; el Etna, circundado por la rompiente siciliana; Chimborazo, en el Ecuador; Mayon, en las Filipinas; Osorno, en Chile, y el pico japonés del Fujiyama, o Fuji, como se lo llama a menudo, cuyo cono se destaca en majestuoso aislamiento por sobre la campiña.

La cumbre de un volcán generalmente exhibe una depresión en forma de cáliz, o de cuenco, que se conoce con el nombre de cráter. Indudablemente, la forma cónica característica de un volcán está causada por el modo en que surge la montaña. Es una acumulación de lava, fragmentos rocosos y escoria proyectados por la chimenea, todo lo cual asume la forma de un cono, con un conducto, o embudo, que atraviesa su centro y llega hasta la cavidad caliciforme de arriba. El ángulo del

D. Cavillon / H. Tazieff

Se elevan nubes de un lago de lava en el cráter central del Nyiragongo, volcán activo que se halla en el Zaire, África, cerca de la frontera de Ruanda.

como depende de la viscosidad, o lentitud, del flujo de lava y de la cohesión del material expelido; cuanto más viscosa es la lava y más cohesivo el material, tanto más escarpado es el cono. Los volcanes de la isla Reunión, en el océano Índico, se formaron a partir de lava muy viscosa y sus laderas caen casi a pique; por el contrario, en las islas Hawai, los conos fueron generados por lava fluida y tienen una pendiente suave. Sus bases pueden medir 100 kilómetros, o más, de diámetro.

La composición de los conos de los diversos volcanes difiere; algunos están constituidos por fragmentos cenicientos o parecidos a escorias; otros, por capas de lava, y aun otros, por una mezcla de ambos. Los volcanes más macizos usualmente están conformados por capas alternadas de lava, fragmentos rocosos y polvo volcánico; cruzando estas capas en varias direcciones hay numerosas grietas llenas de lava. Puede resultar difícil creer que montañas volcánicas de 4.500 ó 6.000 metros de alto se hayan erigido de esta manera, capa sobre capa; sin embargo, no debemos olvidar que este proceso ha estado produciéndose durante muchos miles de años.

FORMA CAMBIANTE

Las montañas surgidas de este modo son susceptibles de desmoronarse así como de erigirse, y pueden alterar considerablemente su forma durante su crecimiento. Las explosiones pueden arrancarles parte de la cima o formar nuevos conos y chimeneas. Con el correr del tiempo, el Vesubio cambió de forma muchas veces; en el siglo I a.C., la cima del volcán era una gran planicie deprimida de casi cinco kilómetros de ancho. Allí, el gladiador Espartaco y sus seguidores fueron sitiados por un ejército romano. Después de la gran erupción del 79, que destruyó las ciudades de Pompeya y Herculano, el Vesubio formó un inmenso cráter, dentro del cual se originó un nuevo cono con un cráter más pequeño.

En 1822, una gran erupción redujo en 120 metros la altura del volcán y produjo un cráter enorme de alrededor de un kilómetro y medio de diámetro y 300 metros de profundidad. Hacia 1843, dentro del gran cráter habían surgido tres conos pequeños con otros tantos cráteres. Las erupciones de 1872 y 1906 alteraron nuevamente la forma del Vesubio. En 1922, dentro del cráter antiguo apareció un nuevo cono de 70 metros de altura. Todos los volcanes activos están sujetos a tales cambios de forma. El efecto general de la actividad volcánica es aumentar tanto el volumen como la altura.

Casi todos los grandes volcanes desarrollan nuevas chimeneas, que dan origen a conos secundarios, o "parásitos", los cuales tienden a destruir la forma cónica original. Los volcanes de las islas Hawai poseen miles de conos secundarios de este tipo; el tamaño de los cráteres varía. Por lo que sabemos, hay escasa vinculación –si es que hay alguna– entre el tamaño del cráter y la altura del volcán. Por ejemplo, el mexicano extinguido de Orizaba tiene una altura de 5.550 metros, con un cráter de alrededor de 300 metros de diámetro. El Popocatepetl no es tan elevado –5.450 metros– y con todo su cráter es casi el doble del de Orizaba. Haleakala, volcán de las islas Hawai, sólo tiene 3.000 metros de altura, y su cráter, 32 kilómetros de circunferencia. No parece haber relación entre el potencial explosivo de un volcán y el tamaño de su cráter.

CALDERAS

Se ha dado este nombre a cráteres muy amplios y comparativamente poco profundos. Por primera vez se denominó así un hoyo muy grande en las islas Canarias: tiene unos 5 kilómetros de diámetro, y lo rodean riscos que se elevan hasta casi 900 metros. Una caldera es el resultado de la explosión, o derrumbe, de un cono volcánico previo, lo cual hace que el cráter se vuelva más amplio y menos profundo.

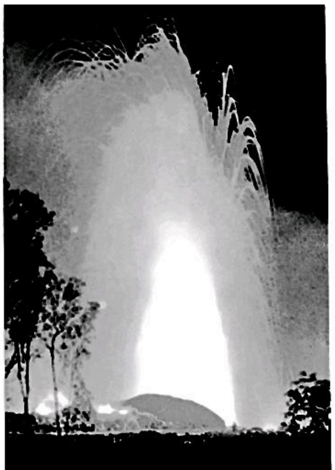
Existen muchas calderas, algunas de ellas enormes; la de Mauna Loa, en la isla Hawai, es de poco más de cinco kilómetros y medio de largo y alrededor de tres de ancho. En la misma isla, la caldera de Kilauea tiene cuatro kilómetros de largo por unos tres de ancho. Ambas se formaron por el derrumbe de los antiguos conos volcánicos.

Derecha: un vulcanólogo mide la temperatura de la lava dentro del cráter del monte Etna, Sicilia, el volcán activo más elevado de Europa. Abajo: vista del Etna durante una de sus violentas erupciones, que emiten inmensas cantidades de vapor de agua.

D. Cavillon / H. Tazieff

Muchas calderas se llenan de agua, y albergan lagos. Entre los más bellos y pintorescos se halla el Crater Lake, en lo que ahora es el Parque Nacional de Crater Lake, en el sudoeste de Oregón, sobre la cresta de las montañas Cascade. Crater Lake está a más de 1.800 metros sobre el nivel del mar; de forma casi circular, lo circundan paredes de piedra de una altura que oscila entre 270 y 670 metros. El lago tiene alrededor de ocho kilómetros de diámetro y, en algunos lugares, cerca de 600 metros de profundidad. Los indios lo llamaron "Mar del Silencio". Para el poeta norteamericano Joaquín Miller, era "un mar de zafiro, circundado por un compacto círculo de... grisáceas rocas".

H. Tazieff



H. Tazieff

El cráter del Kilauea, de Hawai, vomita fragmentos de piedra candente. A menudo se producen terremotos, cuando ocurren las erupciones del Kilauea.

ALGUNOS ACTIVOS, ALGUNOS EXTINGUIDOS

Los volcanes varían mucho en su actividad. Por regla general, períodos de inactividad parcial o completa alternan con otros de explosiva violencia. Hasta la erupción del año 79, el Vesubio había estado en reposo durante siglos. La erupción del Krakatoa aconteció en 1883, después de un período inactivo de doscientos años. Algunos volcanes están activos constantemente; el Stromboli, una isla volcánica del Mediterráneo perteneciente al archipiélago italiano de las Lipari, ha estado vertiendo lava por más de 2.000 años. Izalco, en El Salvador, está en continua erupción desde 1770. Los volcanes de este tipo son relativamente tranquilos.

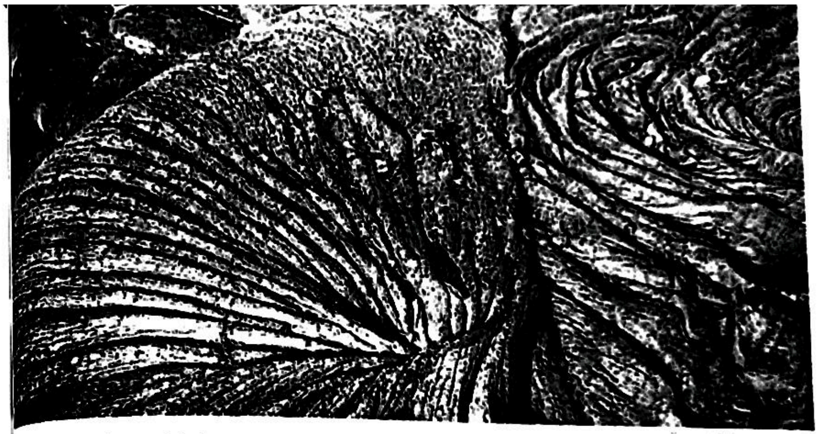
Ciertos volcanes están total o casi totalmente extinguidos; entre éstos se cuentan el gran volcán de Palma, en las Canarias; Monte Shasta, en California; Monte Hood, en Oregón; y Monte Rainier, en Washington. Los picos de ciertos volcanes extinguidos están cubiertos de nieve. Sus laderas han sido esculpidas por la erosión hasta tal punto que van perdiendo rápidamente su forma característica. En algunos casos, las fuerzas erosivas han desgastado antiguos volcanes casi hasta sus cimientos.

VOLCANES DE LODO

En ocasiones, el gas que escapa de una chimenea en la superficie de la tierra puede arrastrar consigo partículas de arena y arcilla. En el trascurso del tiempo se puede formar un cono de tales partículas que, al ser mojadas por fuertes lluvias, se convierten en lodo seco y endurecido en la superficie del montículo; luego, el gas se acumula debajo de esa superficie endurecida, y la presión va haciéndose más fuerte hasta que, al final, el gas hace volar la parte superior del cono. Estos *volcanes de lodo*, como se los denomina, se hallan frecuentemente en campos petrolíferos y gasíferos. Algunos surgen en regiones donde escapa vapor volcánico a través del lodo. Ciertos volcanes de lodo alcanzan alturas considerables: la del Boga-Boga, en la región de Bakú, en la Unión Soviética, supera los 30 metros.

El cráter del Rano Raruku en el sudeste de la Isla de Pascua. Se usaron rocas volcánicas de la ladera de este cráter para hacer algunas de las célebres estatuas de la isla.

Jim Woodman



Lava enfriada proveniente de una erupción en Hawai en 1969. Esta apariencia retorcida y acordonada es característica de las lavas *pahoehoe* muy líquidas, cuando se endurecen

DISTRIBUCIÓN DE LOS VOLCANES

Existen varios miles de volcanes en el mundo, y se sabe que hay alrededor de 450 activos; están distribuidos en una serie de cinturones. El gran cinturón del océano Atlántico incluye un número considerable de islas volcánicas, entre ellas la de Spitsbergen Occidental, la isla Jan Mayen, Islandia, las Azores, Madeira, las islas Canarias y las de Cabo Verde. Otro cinturón volcánico se extiende desde Arabia Saudita hacia el Sur; comprende el mar Rojo, Etiopía y la gran isla de Madagascar, frente a la costa oriental de África.

En los confines orientales del océano Pacífico hay una cadena de volcanes que va desde las islas Aleutianas hasta el cabo de Hornos. Abundan los activos en las cadenas de las Aleutianas y en las costeras de Alaska. El

El Monte Ruang, que se eleva a cerca de 3.300 metros, es uno de los varios volcanes de Java.

Islandia, asentada sobre uno de los cinturones volcánicos de la Tierra, posee muchos géiseres. La foto superior muestra una chimenea, o fisura, de azufre caliente, en el lago Nyvatn, Islandia. Izquierda: un verdadero géiser, un surtidor de vapor de agua, también en Islandia.

volcán Katmai entró en erupción en 1912 y cubrió la isla Kodiak con un manto de lava y ceniza. La cordillera Cascade, en los Estados Unidos, tiene muchos volcanes, pero sólo unos pocos están activos. El pico Lassen, en el norte de California, hizo erupción violentamente en 1914-15. Los montes Rainie, en Washington, y Hood, en Oregón, despiden vapor. En México, desde Colina hasta Tuxtla, hay volcanes activos. El Popocatepetl, en el centro de la nación, despide nubes de humo.

En la América del Sur, una serie de volcanes comienza con el Nevado de Tolima, en Colombia, y continúa hasta el Corcovado, en el sur de Chile. El Cotopaxi, en Ecuador, tiene 5.900 metros de altura; es probablemente el volcán activo más elevado.

Otra línea se extiende desde el estrecho de Bering hasta el límite occidental del Pacífico y luego baja hasta el océano Indico. La península de Kamchatka tiene alrededor de catorce volcanes activos. Las islas del Japón son una larga cordillera volcánica que surgió del océano. La cadena de volcanes pasa por las Filipinas, las Molucas y las islas de la Sonda, y llega hasta Nueva Guinea.

AGREGADO DE MATERIAL SUPERFICIAL

La condensación de las enormes cantidades de vapor acuoso que surge inclusive de volcanes activos pequeños da por resultado el agregado de cantidades considerables de agua a la provisión de la Tierra. Se estima que el monte Etna, en una sola erupción, despidió unos 2.270.000.000 de litros de agua convertidos en vapor. Un pequeño cono parásito del mismo cráter arrojó alrededor de 1.740.000.000 de litros en cien días. Se cree que parte del vapor de agua de las chimeneas volcánicas proviene del agua subterránea que se ha escurrido profundamente en la tierra y ha sido sobrecalentada; y parte proviene de la contenida en las rocas cuando se formaron.

Los volcanes han expulsado masas increíblemente grandes de sólidos del seno de la Tierra. Todos deben su masa al interior de ella. Ya estén extinguidos o activos, ya midan sólo unos 100 metros o casi 10 kilómetros de altura, a ellos se debe la agregación de cantidades enormes de materiales a la superficie de la Tierra. La descarga anual promedio de sedimentos del río Misisipí, en Estados Unidos, es de unos 560.000.000 de toneladas. Durante la erupción de 1902, en unas pocas horas el monte Pelée arrojó quinientas veces esta cantidad en forma de polvo volcánico.

SEÑALES DE ADVERTENCIA

Cuando un volcán está por entrar en erupción, es común que diversos fenómenos adviertan del peligro, como terremotos y fuertes ruidos sordos semejantes a truenos, causados probablemente por el movimiento de gases y roca fundida sometidos a gran presión. También pueden aparecer súbitamente manantiales termales o caldas en la vecindad de un volcán poco antes de su erupción, o quizá se advierta un incremento considerable de su actividad. El nivel de los lagos cercanos puede ascender o bajar apreciablemente, o bien éstos pueden quedar totalmente secos.

LA ERUPCIÓN PROPIAMENTE DICHA

Entonces se produce la erupción propiamente dicha; del cráter surgen cantidades enormes de vapor de agua sobrecalentado, y otros gases que arrastran consigo enormes cantidades de rocas, piedras, cenizas, lava y otros materiales. El suelo tiembla violentamente, y se oye un tremendo rugido; se dice que en 1815, durante la erupción del Tambora, un volcán de la isla Sumbawa, en Indonesia, el rugido podía oírse a una distancia de 1.500 kilómetros.

La columna de gases y partículas sólidas que brotan de un volcán en erupción ascienden verticalmente a gran altura. La fricción de las partículas que salen del conducto crea electricidad estática; aquí y allá destellan relámpagos ramificados y se oyen tronidos. Esta columna puede iluminarse por la lava incandescente del cráter, que entonces parece estar en llamas. En algunos casos, la columna es tan densa que oculta al Sol y crea una oscuridad impenetrable. A medida que se eleva, la condensación del vapor de agua provoca una lluvia torrencial; este fuerte aguacero también está causado parcialmente por la condensación del aire saturado de humedad y elevado a grandes altitudes por la corriente ascendente ocasionada por la salida de los gases volcánicos. Al caer torrencialmente la lluvia, por las laderas del volcán fluyen ríos fangosos que pueden sepultar ciudades al pie de la montaña.

GASES VOLCÁNICOS

La gran mayoría de los gases surgidos de las chimeneas volcánicas está constituido por vapor de agua –por sobrecalentado– que puede alcanzar una temperatura de más de 500° C. Otros gases presentes pueden incluir ácido

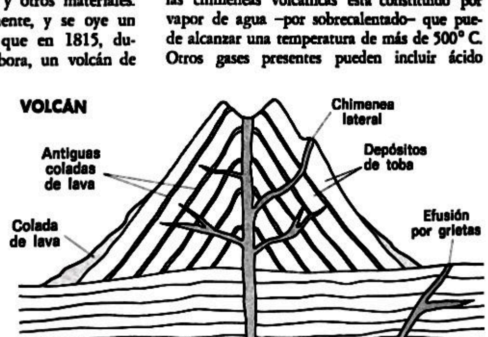


Diagrama de un volcán que muestra cómo el magma –o roca fundida proveniente del interior– pasa a través de capas suprayacentes de rocas, a través de una fisura, o chimenea, para hacer erupción en la superficie, vomitando lava, la cual se acumula en torno de la abertura y forma frecuentemente una montaña. La depresión central en la cima de la montaña volcánica es su cráter.

clorhídrico, bióxido de carbono, monóxido de carbono, metano, hidrógeno, oxígeno y argón, todos a elevada temperatura. Algunos volcanes emiten varios compuestos de azufre, ácido sulf-hídrico y anhidrido sulfuroso. Los gases no sólo surgen del propio cráter volcánico, sino tam-bién de las coladas de lava, que pueden conti-nuar despidiéndolos durante semanas, e incluso meses, después de la erupción.

LAVA

En cierto modo, la lava es el producto más característico de una erupción volcánica. Es cierto que en algunos casos –como en la gran erupción del monte Pelée, en 1902– no se arrojó nada de lava; sin embargo, por regla general, una vez que el volcán "se ha limpiado la garganta", con la descarga de vapor sobre-calentado y otros gases, comienza a fluir lava desde el cráter o, inclusive, éste la arroja hacia lo alto, como un surtidor de metal fundido. Usualmente se derrama por los bordes del crá-ter o mana de fisuras a los lados del volcán; cuando hay muchas de dichas fisuras, el vol-cán parece "sudar" lava por numerosos poros.

CÚPULAS TAPÓN

En algunos casos, cuando la lava despedida es demasiado viscosa para fluir, ya sea por su temperatura comparativamente baja o por su bajo tenor gaseoso, se acumula en forma de grandes cúpulas sobre las chimeneas, y estos edificios volcánicos se denominan *cúpulas ta-pón*. En el Parque Nacional Volcánico Lassen hay trece cúpulas de este tipo en una super-ficie de aproximadamente 130 kilómetros cua-drados.

Un tipo de cúpula tapón fuera de lo común apareció en el monte Pelée tras la desastrosa erupción de 1902. La lava que llenaba el conducto se endureció después de la erupción como roca en su superficie exterior; empero, el interior era muy viscoso, y toda la masa fue gradualmente empujada hacia arriba, hasta que formó una torre de más de 300 metros de altura sobre la cima del volcán. Se la llamó la aguja del monte Pelée. Gradualmente se desplomó, principalmente a causa de explosio-nes de gases dentro de ella. Entre 1929 y 1932 se formó otro tapón, durante una serie de erupciones del monte Pelée.

CÓMO FLUYE LA LAVA

Al principio, el flujo de la lava que surge de la Tierra es a menudo veloz y varía según la viscosidad de la lava y la pendiente del terreno. Cuando alcanza una ladera muy es-carpada se vierte tumultuosamente por ella. Sin embargo, en pendientes suaves puede des-plazarse muy lentamente y, a medida que se enfría, se vuelve más viscosa, y desciende mu-cho más despacio.

El aspecto superficial de la lava enfriada depende, en parte, de su composición química, la que determina si la roca resultante será un basalto, una riolita o una andesita. La natura-leza de la colada de lava es otro factor extre-madamente importante. Las lavas particular-mente viscosas se fragmentan a medida que se enfrían en multitud de bloques escabrosos que se rozan entre sí mientras la lava continúa fluyendo, y a menudo se acumulan en monto-nes o montículos. Cuando finalmente se detie-ne la colada, el resultado es una formación por demás irregular, una acumulación de blo-ques de aristas afiladas y fragmentos mellados. Esta lava áspera se llama *lava en bloques;* a veces se la conoce por el nombre hawaiano de *aa*. La más fluida adopta, cuando se endu-rece, una estructura retorcida y acordonada, con superficie jaspeada y satinada, que recibe el nombre hawaiano de *pahoehoe*.

Cuando acaba de emerger de un cráter, la lava está candente o al rojo blanco; es mala conductora del calor y se enfría muy lenta-mente; en consecuencia, después que se ha formado la corteza superficial y se ha enfriado suficientemente para que se pueda caminar so-bre ella, aún puede estar candente por debajo. Se podrían dar muchos ejemplos de la lenti-tud con que se enfría. En 1830, todavía se-guía saliendo vapor de agua de lava que había fluido del Etna, en Sicilia, cuarenta y tres años antes; el volcán de Jorullo, en México, des-cargó ríos de lava en 1759; ochenta y siete años después todavía se elevaban del lugar dos columnas de vapor.

ROCAS VOLCÁNICAS

Además de lava, tal como hemos visto, los volcanes descargan enormes cantidades de ro-cas, piedras, escorias y cenizas. Durante la erupción de 1779, el poderoso Vesubio arrojó escorias hasta una altura de 3.000 metros. En 1815, el Tambora, ya mencionado, cubrió el mar en muchos kilómetros con tal enorme cantidad de piedra pómez, que los buques ape-nas si podían navegar a través de ella. Se dice que un volcán en el norte de Ecuador central, el Cotopaxi, despidió un bloque de doscientas toneladas a catorce kilómetros y medio.

La acumulación de lava, rocas y arena en la vecindad de un cráter alcanza a veces pro-porciones casi increíbles. La creación del vol-cán conocido como el Paricutín es un ejemplo notable.

El 20 de febrero de 1943 hubo un fuerte terremoto en el estado de Michoacán, México, a unos 290 kilómetros al oeste de la Ciudad de México. Después del terremoto apareció un cráter en un maizal y comenzó a arrojar pie-dras y arena; luego vomitó inmensos ríos de lava. La acumulación de materiales volcánicos en torno del cráter crecía a velocidad fantástica; pronto nació un nuevo volcán, el Paricutín. En el lapso de un año, su altura llegó a ser más de un tercio de la del Vesubio, cuya forma-ción tardó muchos miles de años.

POLVO Y VAPOR DE AGUA

Los volcanes arrojan una vasta cantidad de polvo. Cuando el Tambora hizo erupción en 1815, cayó abundante cantidad de aquél sobre Borneo, a 1.400 kilómetros de distancia. Du-rante la erupción del Cosigüina –o Consegüi-na–, en Nicaragua, se estima que este volcán arrojó a la atmósfera unos 4.740.000.000 de metros cúbicos de polvo, que los vientos arras-traron 1.300 kilómetros por lo menos. En 1877, el polvo proveniente de una erupción del Cotopaxi sumergió a la ciudad de Quito, Ecuador, en la más absoluta oscuridad.

El polvo de la erupción del Krakatoa, en 1883, oscureció el cielo hasta 240 kilómetros del volcán y cayó en cantidades apreciables a 1.500 kilómetros.

El polvo en las proximidades de un volcán en erupción ocasiona una oscuridad pavorosa, sólo interrumpida por relámpagos de luz pro-venientes del cráter y por la incandescencia y el resplandor de la lava fundida; luego, produce efectos de pasmosa belleza en el cielo. A medida que flota hacia las alturas –circun-dando varias veces el globo en ciertas oca-siones– el fino polvo volcánico refracta o re-fleja cada uno de los colores de la luz solar, produciendo así amaneceres y ocasos magní-ficos en todo el mundo.

El diluvio proveniente del vapor condensa-do de la erupción suele convertir el polvo vol-cánico en un fango pegajoso que arrolla y cubre cuanto esté en su camino. La antigua ciudad romana de Herculano fue sepultada durante la erupción del Vesubio, en el año 79 de nuestra era, por una inundación de lodo. La gran erupción del Cotopaxi, en 1877, cubrió muchos pueblos con barro y lava y varias clases de detritos. Bosques enteros fueron abatidos como consecuencia de este fenómeno.

Tony Gauba. Audubon PR.

Lava demasiado viscosa para fluir se acumula en grandes cúpulas, o torres, y forma un cuello, o tapón volcánico. He aquí un cuello vocánico conocido como Shiprock, al noroeste de Nuevo México.

VOLCANES SUBMARINOS

Los volcanes submarinos figuran entre los más espectaculares: el Etna, Stromboli, Pico de Tenerife y muchos otros tuvieron origen bajo el mar. Nadie sabe cuántos volcanes activos hay ahora en las profundidades oceánicas. A menudo arrojan escorias y vapor y agitan la superficie marítima; estas erupciones vienen acompañadas de vastas efusiones de lava sobre el suelo oceánico; a veces forman varias capas superpuestas y, finalmente, sobre la superficie del océano aparece una isla. La cadena volcánica de las islas Hawai se formó de esta manera.

NACIMIENTO DE LAS ISLAS VOLCÁNICAS

En ciertas ocasiones, los volcanes submarinos descubren su presencia del modo más sorprendente: de pronto puede emerger una isla nueva, o desaparecer una antigua bajo la superficie. Esta acción es muy diferente de la gradual expulsión de lava.

El mar Mediterráneo ha sido a menudo testigo de la súbita aparición y desaparición de nuevas islas volcánicas. En 1831, repentinamente surgió, entre Sicilia y África, una nueva isleta –llamada isla de Graham– como resultado de la acción volcánica. Alcanzó una altura de 60 metros sobre el nivel del mar y una circunferencia de alrededor de cinco kilómetros; como estaba compuesta por materiales sueltos, no pudo resistir el embate constante del mar: fue rápidamente erosionada por las olas y, en unos pocos meses, quedó reducida a un bajío, llamado arrecife de Graham. Frecuentemente surgen nuevas islas entre Sicilia y Grecia, pero en general desaparecen casi tan rápidamente como aparecieron.

Algunas islas producidas por volcanes submarinos tienen mayor permanencia, como tres del mar de Bering; una, llamada isla Bogoslov, apareció a unos 65 kilómetros al oeste de la isla Unalaska, en 1796, después de una erupción. En 1883, otra erupción proyectó un cono volcánico de arena negra y cenizas, conocido como Nueva Bogoslov, o isla de Fuego. Todavía otra más hizo su aparición en 1906.

En 1963 surgió, frente a la costa suroccidental de Islandia, cerca de las islas Vestmannaeyjar, otra volcánica, llamada después Surtsey; en la actualidad, dicha isla alcanza cerca de dos y medio kilómetros cuadrados y más de 150 metros de altura.

ERUPCIONES DE FISURA

No todos los materiales de origen volcánico se acumulan para formar montañas; en muchos casos, la lava fluye de grietas en la corteza terrestre y se extiende por grandes superficies sin producir nada que sugiera conos volcánicos. A tales derrames se los denomina *erupciones de fisura*. El geólogo norteamericano H. S. Washington propuso que se los llamara flujos de meseta, pues la superficie que cubren es plana como ésta. La lava que emana de erupciones de fisura, o flujos de meseta, no es muy viscosa.

Las erupciones de fisura se han producido en vasta escala. En la región del Decán, en la India occidental, una zona de más de 500.000 kilómetros cuadrados ha sido cubierta de lava, que alcanza en algunos lugares un espesor de 600 metros. Partes de Washington, Oregón, Idaho, Montana y California yacen bajo un mar de lava solidificada de alrededor de 575.000 kilómetros cuadrados de superficie, con un espesor medio de 150 metros. Flujos similares han cubierto grandes regiones de la parte norte de Wisconsin y Michigan y la costa norte del lago Superior. En un tiempo, unos 100.000 kilómetros cuadrados de Inglaterra estuvieron bajo un mar de lava de fisura. Grandes comarcas de Etiopía se encuentran cubiertas por este tipo de lava.

Las erupciones de fisura han sido comunes en Islandia, isla erigida en gran parte por la actividad volcánica. En 1783, de una fisura surgieron dos corrientes de lava que fluyeron 65 y 45 kilómetros, respectivamente. El gran desierto de lava de Odahahraun en Islandia –que cubre una superficie de más de 3.500 kilómetros cuadrados– también es producto de erupciones de fisura.

POR QUÉ OCURREN LAS ERUPCIONES

Las erupciones volcánicas y las efusiones de lava a través de fisuras se deben al calor interno de la Tierra. No se sabe con certeza el origen de éste. Algunos geólogos sostienen que es un residuo del estado de fusión original de la Tierra durante el período inicial de su formación. Dicen que su calor interno ha sido conservado por tantos millones de años porque las rocas son muy malas conductoras. Otros sostienen que el calor se produce por la descomposición de elementos radiactivos, tales como uranio y torio, ambos presentes en todas las rocas, si bien en cantidades generalmente muy pequeñas.

Cualquiera sea su origen, el calor terrestre es suficientemente intenso para licuar grandes masas rocosas. Este material fundido, llamado *magma*, consiste principalmente en soluciones de silicatos con óxidos y sulfuros. También contiene vapor de agua y otros gases, retenidos en solución por la presión. Allí donde se forma este magma, la roca líquida pesa menos que la sólida adyacente, y los gases disueltos en ella la hacen todavía más liviana.

El magma está sometido a la presión de la roca más pesada que lo circunda y se ve forzado hacia arriba. Parte de él alcanza la capa externa de la corteza terrestre y se abre camino hacia la superficie a través de fisuras o chimeneas. Luego, las fuerzas explosivas del interior de la tierra lanzan el material fundido hacia lo alto, o bien éste fluye a través de cráteres o de grietas en forma de lava más o menos viscosa; la lava es simplemente magma que ha alcanzado la superficie de la Tierra y que ha perdido la mayor parte de sus gases.

GÉISERES

Los géiseres son chorros de agua hirviente que salen de la tierra en unas pocas regiones volcánicas. El término deriva de la palabra islandesa "geysir", que significa "estallar con violencia". El verdadero géiser es un brillante surtidor de vapor de agua que, a menudo, se eleva a varios cientos de metros en el aire, con una persistencia que dura desde unos pocos minutos hasta una hora o más. Mana a borbotones en una zona de extensos manantiales termales. Sólo hay tres regiones principales donde existen géiseres: Islandia, Nueva Zelanda y el Parque Nacional del Yellowstone, en Estados Unidos; de las tres, la del Yellowstone es la que posee el más espectacular despliegue de géiseres.

YELLOWSTONE

El fantástico espectáculo que ofrecen los chorros de vapor de agua del Yellowstone probablemente fue vislumbrado inicialmente por una partida de exploración oficial en 1807, en

Steve McCutcheon, Servicio Gráfico de Alaska

El lago Emerald se ha formado en el cráter del monte volcánico Katmain, en Alaska. La nieve cubre la región circundante.

la época de la Expedición de Lewis y Clark al Territorio del Noroeste. Luego, tramperos y batidores exploraron el Yellowstone y a veces enviaban informes a Washington. Pese al cúmulo de datos, los habitantes de la región oriental permanecían escépticos respecto a la existencia de los géiseres. En 1870, Nathaniel Langford publicó en *Scribner's Monthly* una narración de un viaje de varias semanas a través de la región del Yellowstone. La sensacional descripción de Langford sobre el valle Firehold, con sus 3.000 fuentes hirvientes, sus cien géiseres turbulentos y sus muchas fuentes deslumbrantes de vapor de agua, fue rechazada despectivamente y considerada por los expertos de salón en el este como un "disparate". Langford mismo fue denunciado como el "campeón de los mentirosos del Noroeste". Cuando David Folsom, que realizó un viaje similar en 1869, intentó publicar sus experiencias, recibió una nota sumamente irónica del *Lippincott's Magazine*: "Gracias, pero no publicamos ficción."

De todos modos, el interés por la singular maravilla natural del Yellowstone continuó. El informe oficial de Gustavus C. Doane, subteniente del 2º Cuerpo de Caballería de los Estados Unidos, miembro él mismo de la partida Langford, probó ser el estudio más completo realizado hasta entonces sobre el Yellowstone. Comenzó un movimiento para incorporar el territorio a los dominios nacionales. No se pudo elegir un momento peor; el país se ha-

llaba atenazado por los escándalos de la Administración Grant, y había una fuerte presión en el Congreso en favor de entregar el Yellowstone a la explotación privada. Gracias al gobierno de los E.U., la zona se salvó, en beneficio permanente de todos.

El 1º de marzo de 1872, el Congreso aprobó el Acta que creaba el Parque Nacional del Yellowstone como "terreno de esparcimiento público y reserva de caza". Nathaniel Langford fue nombrado primer superintendente de este magnífico santuario de vida silvestre. En la actualidad, el Parque Nacional del Yellowstone posee una superficie total de 8.850 kilómetros cuadrados y tiene casi tres veces la extensión del Estado de Rhode Island.

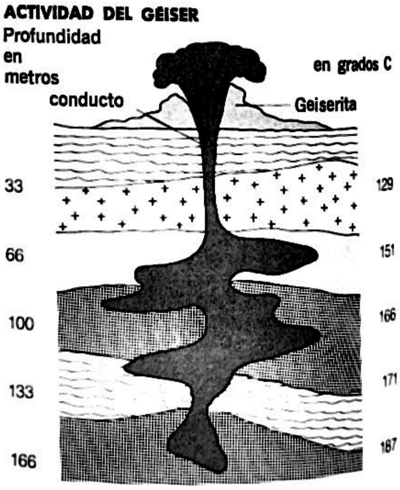
Durante la década de 1870, un visitante inglés, el cuarto conde de Dunraven, viajó a través del Yellowstone a la búsqueda de caza mayor, con Buffalo Bill y Texas Jack. Dunraven quedó fascinado ante el valle Firehole y la furia de los géiseres: la Giganta, el Castillo, el Viejo Fiel y el Gran Géiser; este último lanzó una columna de vapor a casi 30 metros de altura, desde cuyo ápice surgían cinco chorros torrenciales hasta una altura de alrededor de 75 metros del suelo.

En *The Great Divide*, publicado en 1874, Dunraven describió el Géiser del Castillo con las siguientes palabras: "Pudimos oir un gran ruido... La perturbación fue más y más sonora, hasta que, en un súbito estertor, surgieron unas pocas toneladas de agua, y así obtuvo el géiser alivio momentáneo. Después que unas pocas expulsiones premonitorias nos previnieron que retrocediéramos una corta distancia, los síntomas empeoraron rápidamente; el rugido y la baraúnda aumentaron su intensidad; la agonía del monstruo se volvió cada vez más violenta; la tierra temblaba ante su furia y, finalmente, con un poderoso espasmo, lanzó al aire una gran columna de agua que, diría yo, alcanzó en su punto más alto una altura de 250 pies [76 metros]. Lo rociada y el vapor arrojados alcanzaron mayor altura aún y luego flotaron como una nube densa a una altitud indefinida. La operación no era continua, sino que consistía en una pulsación fuerte y clara que se producía a razón de un máximo de setenta por minuto, con tendencia a aumentar gradualmente su fuerza y rapidez, hasta alcanzar el máximo desarrollo de intensidad para luego disminuir en forma gradual. Pero el incremento y la disminución no eran uniformes ni regulares; surgían los chorros, que se volvían cada vez más fuertes con cada pulsación durante diez o doce efusiones, hasta que el esfuerzo culminaba en tres impulsos de insólita energía... El volumen de agua proyectado debe haber sido prodigioso; la rociada descendía en forma de lluvia intensa cubriendo gran superficie, y descendían torrentes de agua caliente de seis u ocho pulgadas [15 ó 20 centímetros] de profundidad, por la plataforma inclinada."

En la Cuenca Media de la región del Firehole está el "Tarro de Pintura del Diablo", una caldera de lodo hirviente con muchas tonalidades, que burbujea y despide vapor en un juego maravilloso de colores cambiantes.

El agua caliente arrojada por los géiseres siempre es hermosamente clara, pero, sin embargo, tiene un alto tenor de materia mineral, principalmente sílice disuelta. En torno de los orificios de los géiseres, las incrustaciones de esos materiales toman con frecuencia formas y colores hermosos o fantásticos; así, el Géiser del Castillo se ha vuelto un bello cono blanco, y el Géiser de la Gran Fuente, un pedestal ancho y circular de alrededor de 60 centímetros de altura. Las algas que crecen exuberantes en el agua caliente producen los tintes que presentan los depósitos minerales.

Los géiseres hacen erupción cuando el agua subterránea se pone en contacto con el calor volcánico en una fisura vertical y alcanza su punto de ebullición.



GÉISERES ISLANDESES

Si bien el Parque del Yellowstone se enorgullece de poseer la mayor cantidad de géiseres y los más famosos, no se debe dejar de lado a Islandia, lugar de origen del estudio de este fenómeno. En sí mismo, este país es una de las maravillas del mundo; es una tierra de fuego y hielo, de glaciares y géiseres, de lava y agua de deshielo, de aludes y de volcanes.

Desde la cima del Odahahraun se extiende un campo de lava de más de 3.650 kilómetros cuadrados. En total, casi 12.000 kilómetros cuadrados de la isla son negros a causa de la lava, y 13.340, blancos por la nieve. En este desierto blanco y negro surgen grandes volcanes, como el Hekla (1.560 metros). La mayoría de ellos está ahora inactiva, pero el Askja –cuyo cráter tiene una superficie de 87 kilómetros cuadrados y 900 metros de profundidad– estuvo en erupción en 1875 y cubrió con sus cenizas más de 5.000 kilómetros cuadrados.

El Gran Géiser ha producido un montículo de material silíceo de alrededor de 12 metros de altura; y desde la cuenca en forma de platillo de su cima desciende por unos 23 metros un tubo de tres metros de diámetro. En sus comienzos solía lanzar agua hasta una altura de 45 metros, pero ahora la altura promedio de su chorro es de sólo 20 ó 25. El ruido del géiser en erupción se ha comparado al rugir del mar furioso, entremezclado con disparos de armas que se repiten a intervalos regulares.

El Strokkr no tiene una cuenca regular: es simplemente un tubo con forma de embudo que se estrecha desde un diámetro de 2,5 metros, en la superficie, hasta 25 centímetros, 13 metros más abajo. Usualmente se lo puede hacer entrar en actividad arrojándole adentro turba o piedras; la cantidad echada regula el tiempo de erupción. El Strokkr puede expeler agua más alto que el Gran Géiser, pero su altura usual sólo es de 9 a 12 metros.

Respiraderos, o cráteres diminutos, formados en la superficie de depósitos espesos de lava; se hallan a lo largo de la costa de Tongatapu, isla del Pacífico Sur.

GÉISERES DE NUEVA ZELANDA

Una tercera tierra de maravillas, en cuanto a géiseres, es la que se halla en la Isla del Norte de Nueva Zelanda, en el territorio volcánico que circunda el lago Taupo y Rotorua. Aquí encontramos el géiser más grande del mundo. Su tubo, de 24 metros de profundidad, está situado en medio de un lago caliente, ubicado en un cráter que se formó durante la gran erupción volcánica de 1886. Este géiser tremendo no tiene la belleza de los del Parque del Yellowstone, pues la columna de agua que arroja es barrosa y de color negro tinta, pero sobrepasa a los del Yellowstone en violencia. A menudo, arroja su columna de agua hasta una altura de 150 metros y, en una ocasión, alcanzó al triple de esa altura. A veces está inactivo durante semanas y otras entra en actividad, también durante semanas. Cerca de Rotorua existen varios géiseres, pero sólo los más pequeños están activos actualmente. El géiser Wairoa, que acostumbraba brotar espontáneamente hasta una altura de 60 metros, sólo actúa ahora cuando se lo alimenta con barras de jabón.

Se ha descubierto que el agregado de ciertas sustancias alcalinas –tales como jabón o lejía– torna el agua algo viscosa, determinando que retenga las burbujas de vapor más pequeñas hasta que el volumen de vapor de agua acumulado irrumpe con creciente poder eruptivo y un súbito descenso de presión. Antes de la gran erupción de 1886, dos lagos de géiser habían producido, por su desbordamiento, un conjunto de incrustaciones silíceas maravillosas conocidas como la Terraza Blanca y la Terraza Rosa. Estas bellas formaciones fueron destruidas durante la erupción.

MOTIVO DE LAS ERUPCIONES

¿Cuál es la causa de la erupción de los géiseres? Robert Wilhelm Bunsen, famoso químico alemán del siglo XIX, elaboró una teoría muy satisfactoria para explicar el fenómeno. Dicha teoría es aún casi generalmente aceptada.

Se basa en que el punto de ebullición del agua se eleva a medida que aumenta la presión, y desciende a medida que ésta disminuye. El agua hierve a una temperatura de 100° C al nivel del mar, donde la presión atmosférica es de alrededor de 1 kilogramo por centímetro cuadrado. La presión es doble a una profundidad de diez metros, donde el agua hierve a una temperatura de 120 ° C.

Si el agua subterránea penetra en una fisura más o menos vertical y muy profunda, formará una columna líquida a medida que se acumula. Como el fondo de la columna se pone en contacto con el calor volcánico, su temperatura se elevará considerablemente por encima de los 100° C que, como hemos visto, representa el punto de ebullición del agua al nivel del mar. Empero, no ha de hervir al principio, porque está sometida a la presión ejercida por la columna de agua que tiene encima, y su punto de ebullición se eleva.

A medida que el agua del fondo de la columna se vuelve más caliente, parte se convertirá en vapor, cuya presión hará que ascienda la columna de agua, y algo de ella desborde en la superficie; esto significa que menos agua hará presión sobre la del fondo, la presión disminuirá, y descenderá el punto de ebullición. A consecuencia de esto, más agua se convertirá en vapor. Finalmente, el vapor acumulado arrojará fuera de la fisura toda la columna acuosa, y entonces se producirá la erupción de un géiser. El fenómeno se repetirá en forma más o menos regular, según la provisión de agua y calor.

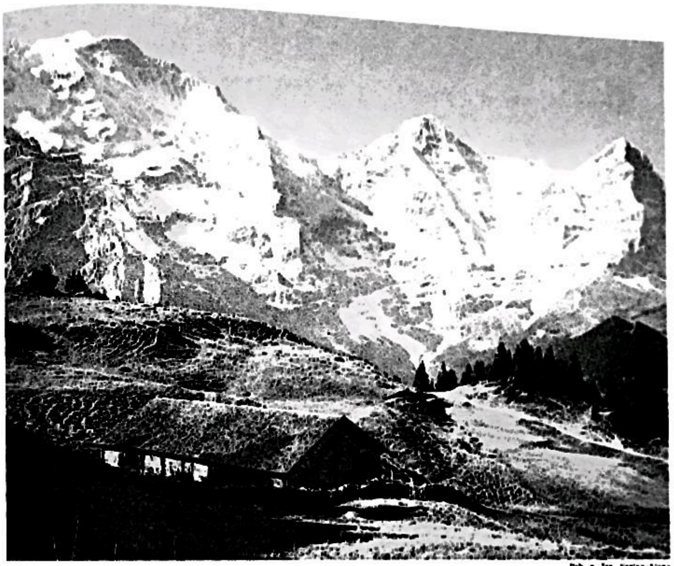
¿PODEMOS UTILIZAR LA ENERGÍA?

La formidable energía térmica de la Tierra, que es fuente de toda actividad volcánica y de géiseres, se conoce como ***energía geotérmica***. El hombre ha comenzado a controlarla y a usarla para sus propias necesidades.

El calor proveniente de fuentes subterráneas se usa ampliamente en Islandia y hasta cierto punto en Nueva Zelanda, Japón y la Unión Soviética, y en algunos lugares de los Estados Unidos. Sin embargo, la energía geotérmica está recibiendo mayor atención como fuente de combustible para la generación de electricidad. Básicamente, el vapor proveniente de un campo geotérmico se emplea para impulsar una turbina que acciona un generador y produce electricidad.

La aplicación comercial de la energía geotérmica para generar electricidad se inició a comienzos del siglo XX en Italia. Hoy en día, en Nueva Zelanda, Japón, los Estados Unidos, la Unión Soviética, Italia e Islandia se produce energía eléctrica a partir de fuentes geotérmicas. El más grande de los campos geotérmicos conocidos es el de Los Géiseres, en el norte de San Francisco, California.

Para un tratamiento más detallado de esta fuente de energía, véase el artículo "Energía geotérmica", en esta misma obra.



Bob e Ira Spring Alpes

Con una duración de millones de años, el proceso imperceptible que formó los Alpes comenzó con sedimentación bajo el mar y luego levantamiento y erosión.

LAS MONTAÑAS

La poética frase "las colinas eternas" es perfectamente válida si comparamos éstas con la breve vida humana. No obstante, contemplada desde el punto de vista del tiempo geológico, la cadena de montañas más antigua sólo dura lo que un simple sueño de una noche en el término de una vida humana. Desde las épocas precámbricas hasta el presente, el perpetuo proceso de construir y destruir montañas ha continuado. Las macizas cadenas llamadas cordilleras, como los dilatados Andes, pueden ser el producto de millones de años de depósitos sedimentarios y plegamientos de la corteza terrestre. Por otra parte, el cono volcánico del Kilimanjaro, en Tanzania, probablemente nació de modo relativamente súbito, como ocurrió con el volcán Paricutín, en México, que se formó en 1943.

COMO SURGEN LAS MONTAÑAS

Los geólogos llaman ***orogenia*** a la formación de montañas, palabra procedente de dos raíces griegas que significan "origen de las montañas". Existen tres clases diferentes de procesos. En el volcánico, se alzan picos aislados, conforme la presión del calor interno de la Tierra entra en juego. La acción volcánica se produce más frecuentemente en las profundidades oceánicas que en tierra. En las escondidas depresiones del mar hay numerosos volcanes submarinos en actividad y ocasionalmente alguno aparece en la superficie. El segundo proceso es la denudación y erosión mediante las cuales elevadas mesetas van convirtiéndose gradualmente en picos y cordilleras por el efecto demoledor de los elementos y la acción erosiva de los ríos. Un

buen ejemplo de meseta original, ahora desgastada en forma de montañas, es la cadena Catskill en el sur del estado de Nueva York. El Gran Cañón del Colorado revela los efectos de la erosión en la meseta existente sobre el río Colorado. En pocos millones de años, toda el área habrá quedado trasformada por el río en picos dentados y grietas.

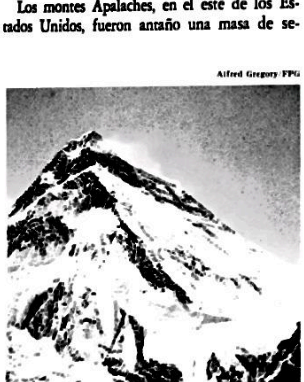
El tercer modo en que surgen las montañas es mediante el depósito de sucesivas capas de sedimentos, ya sea en el fondo de un mar poco profundo, ya en una amplia depresión de la tierra conocida como *geosinclinal.* En el curso de sucesivas épocas geológicas este sedimento se acumula en tal medida que en la corteza terrestre se crea una grave tensión. Otros factores, tales como la presión del calor en el interior de la Tierra y la presencia de elementos radiactivos en la roca circundante, pueden también causar el gradual solevantamiento del depósito sedimentario. El proceso es largo, y probablemente períodos de pausado ascenso alternan con súbitos levantamientos.

Los montes Apalaches, en el este de los Estados Unidos, fueron antaño una masa de sedimentos geosinclinales, tal como hemos descrito anteriormente. Hace más de 500.000.000 de años, el actual emplazamiento de estas montañas estuvo ocupado por una estrecha depresión o geosinclinal, colmada por el agua del mar. Inmediatamente al este de la depresión había una antigua masa de tierra que los geólogos han denominado Apalachia. La opinión presente sostiene que la Apalachia era probablemente una serie de islas que formaban un arco por fuera de la costa oriental de América del Norte. Se supone que la Apalachia proporcionó la mayor parte de los sedimentos al mar que se extendía entre el arco de islas y el gran continente hacia el oeste.

A medida que la depresión se llenaba con los detritos procedentes de la erosión de las montañas y planicies de la Apalachia, sus sedimentos se convertían en rocas. Después de unos pocos episodios menores de formación de montañas, poderosas fuerzas internas de la Tierra tendieron a contraer aquellos lechos rocosos, arrugándolos, plegándolos y haciendo que gradualmente se fueran elevando hasta convertirse en montañas replegadas, no una, sino varias veces. La erosión deprimió los picos, después de cada elevación. Los actuales Apalaches son restos desgastados de cadenas de montañas que otrora podrían haber rivalizado en esplendor con los Himalaya.

Las pruebas respecto a estos procesos de formación de montañas son concluyentes. Las montañas replegadas se componen primordialmente de rocas sedimentarias estratificadas que frecuentemente contienen restos fosilizados de plantas y animales marinos. Luego de muchos intentos infructuosos para explicar de otra manera estos fósiles, los científicos han llegado gradualmente a aceptarlos como prueba del origen de ciertas montañas, o de sus restos, a partir de rocas formadas bajo el agua. Las montañas mismas no tienen necesariamente que haber surgido de las profundidades del mar, el cual podría haber desaparecido de dicha región antes que el proceso orogénico se acelerara.

Alfred Gregory/FPG



El monte Everest, el pico montañoso más alto del mundo, es de desarrollo reciente en la escala de tiempo geológico.

MATERIA PRIMA

Los sedimentos pueden considerarse como la materia prima de la orogenia. Son fragmentos de roca y tierra que los agentes atmosféricos desgastaron de suelos y montañas anteriores, y fueron arrastrados por los ríos al mar, donde lentamente se acumulan capa sobre capa. Cuando se juntan varios kilómetros de sedimentos y, superpuestos, constituyen una masa de muchos kilómetros de longitud, en gran parte ya solidificada en roca, comienza el proceso principal de formación de montañas. Al principio es intermitente, pero al final alcanza un gran clímax de plegamientos, levantamientos y fallas, o agrietamientos.

¿De dónde proceden todos estos sedimentos? Para responder a esta pregunta consideremos los procesos de erosión en la actualidad. Como ejemplo digamos que, anualmente, los ríos Ganges y Brahmaputra de la India y el río Misisipí de Estados Unidos contribuyen juntos con casi mil millones de toneladas de sedimento a los océanos. Aumentemos esta cifra con la cantidad adicional de sedimento trasportada por los otros grandes ríos del mundo y multipliquémosla por cientos de millones de años. Tendremos entonces alguna idea de las impresionantes cantidades de arena y lodo implicadas.

Como las propias montañas, incluso las derivadas de depósitos sedimentarios anteriores, son una importante fuente de dicho material, no es sorprendente que la provisión de sedimento parezca ser casi ilimitada. Aun hoy día se estima que las rocas de la región de los Apalaches tienen en algunos lugares varios kilómetros de espesor. El origen fundamental de esta roca sedimentaria es el aparentemente inextinguible depósito del interior de la Tierra, que de vez en cuando emite vastas cantidades de roca fundida. Luego ésta se solidifica y posteriormente sufre la erosión.

Las mismas montañas plegadas, sin embargo, no dan un verdadero indicio de las inmensas masas de sedimento que contribuyeron a su creación. Cuando los sedimentos se convierten en roca se produce una gran contracción en su volumen. Después, esas rocas son plegadas y comprimidas en un espacio aún más pequeño. Recordemos también que incluso las montañas existentes han sufrido considerable erosión.

Los geólogos han intentado reconstruir la extensión original de los depósitos reuniendo algunos de los principales sistemas de montañas del mundo. En casos extremos, sólo los repliegues y fallas representan ya una mensurable reducción de la corteza terrestre. Pero incluso hoy encontramos rocas antiguas que apenas han sido perturbadas, probando que mucha cantidad de sedimento nunca se ha visto envuelta en procesos orogénicos. Esto revela la superabundancia de sedimento disponible.

Dr. E. R. Degginger

Sedimentos plegados y comprimidos forman la cordillera de Alaska, hogar del monte McKinley.

CRESTAS Y DEPRESIONES

Las montañas no sólo surgieron en el fondo de mares desaparecidos, sino que frecuentemen-

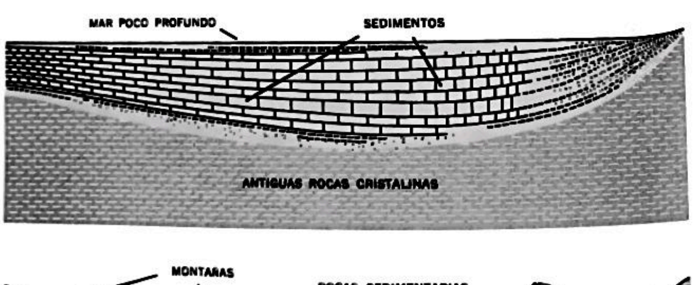
te permanecieron sumergidas hasta mucho después de su formación y luego han vuelto a elevarse. El registro de su sumersión puede leerse en los depósitos que sedimentaron sobre ellas mientras estuvieron debajo del agua. Estos últimos depósitos pueden yacer a varios ángulos respecto del lecho original de las montañas, puesto que se formaron sobre las laderas y crestas.

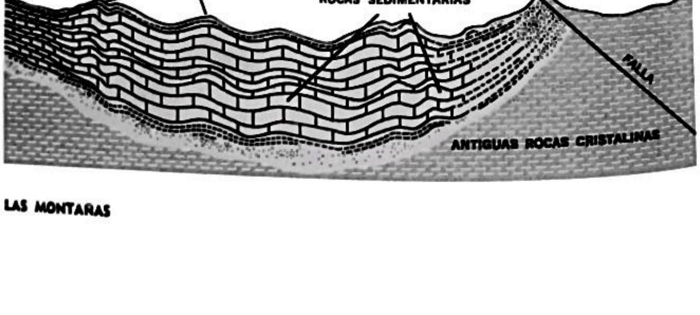
¿Cómo podemos explicar o comprender las poderosas fuerzas que replegaron y alzaron masas gigantescas de roca hasta formar cordilleras? Los geólogos han ofrecido varias explicaciones teóricas, de algunas de las cuales trataremos más adelante en este artículo. Las fuerzas orogénicas actúan muy lenta e imperceptiblemente durante largos períodos de tiempo geológico.

Originariamente, cuando las capas de sedimento se depositaban en el fondo del mar, debieron ser fundamentalmente chatas y horizontales. Sin embargo, si examinamos los principales sistemas de montañas, tales como los Apalaches, las Rocosas o los Andes, descubrimos que los estratos o capas que las integran a menudo no son realmente horizontales. Las capas han sido frecuentemente contraídas en una serie de pliegues. Éstos tienen forma de crestas y depresiones que semejan grandes olas oceánicas. Las crestas o arcos de los pliegues se llaman *anticlinales*. Las depresiones o simas se conocen por *sinclinales*.

Un sinclinal es mucho más pequeño que un *geosinclinal*. Éste es una vasta concavidad de la corteza terrestre que se extiende por cientos o incluso miles de kilómetros a lo largo del eje del pliegue. Por otra parte, un sinclinal puede medir sólo un metro de ancho (aunque pueda ser mucho mayor en longitud). Un geosinclinal puede contener cientos de sinclinales y anticlinales.

Cortes trasversales simplificados de un geosinclinal, o depresión, de sedimentos (parte superior) y de una cordillera formada luego a partir de él (parte inferior). La anchura del geosinclinal y de las montañas se mide de izquierda a derecha; la longitud o eje del geosinclinal se mide en ángulo recto con la página (aquí indeterminados). La escala vertical está muy exagerada. En el geosinclinal, el rayado en forma de ladrillos significa depósitos de arcilla; en las montañas, piedra caliza. Las líneas de trazos representan lodo en el geosinclinal y roca arcillosa (pizarra) en las montañas. El punteado indica arena en el geosinclinal, piedra arenisca en las montañas. Las fuentes de estos sedimentos y rocas son las masas de tierra hacia derecha e izquierda del geosinclinal.





Existen también grandes convexidades de rocas llamadas *geanticlinales*. Un geanticlinal puede hallarse paralelo a un geosinclinal y proporcionarle sedimentos. Tal como el geosinclinal, un geanticlinal puede también estar formado por muchos anticlinales y sinclinales más pequeños.

FALLAS

Los lechos geosinclinales frecuentemente están quebrados o agrietados. Si las rocas a ambos lados de una grieta se hallan desplazadas unas respecto de otras, la grieta recibe el nombre de *falla*. Fallas gigantescas pueden extenderse por varios cientos o incluso miles de kilómetros. Representan zonas donde enormes masas de rocas han sido probablemente desplazadas o apartadas por otros tantos kilómetros. Los repentinos movimientos de las rocas a lo largo de las fallas causan frecuentemente graves terremotos.

Los pliegues y las fallas prueban que formidables fuerzas terrestres actúan o han actuado sobre las rocas de un geosinclinal, trasformándolas en montañas. Las fuerzas pueden haber sido suficientemente poderosas para invertir algunos de los lechos rocosos. En muchos casos, los sedimentos y las rocas sedimentarias que se formaron de ellos han sufrido grandes trasformaciones debidas a la compresión originada por fuerzas poderosas.

Estas fuerzas se ejercían no tanto de arriba abajo como en otras direcciones. La naturaleza de los movimientos terrestres demuestra que las fuerzas horizontales actuaban paralelamente a los lechos rocosos. Sin embargo, las fuerzas verticales produjeron cierta cantidad de pliegues y elevaciones.

TEORÍAS ACERCA DE LA FORMACIÓN DE LAS MONTAÑAS

Una explicación comúnmente aceptada en otro tiempo respecto al origen de las montañas de plegamiento fue la *hipótesis de la contracción*. Ésta sostenía que la Tierra ha venido contrayéndose constantemente, ya sea por causa de la enorme presión de las capas superiores de roca, o debido a un progresivo enfriamiento de un estado anterior candente o fundido. La contracción de la Tierra produciría el arrugamiento

Los estratos de esta roca se han conformado como una serie de pliegues que recuerdan los de grandes olas oceánicas, con crestas y valles claramente visibles. Este es un ejemplo de los tipos de formación conocidos como sinclinal y anticlinal.

Jerome Wyckoff

Cuando las rocas soportan la presion de las masas de tierra adyacentes, se comban y pliegan. Arriba: Un risco plegado en el Parque Nacional del Glaciar.

de toda la corteza. Las arrugas serían los valles y las montañas, y sugieren la idea de los pliegues que aparecen en la piel de una manzana que haya estado secándose y contrayéndose durante algún tiempo.

A comienzos del siglo XX, el astrónomo inglés George H. Darwin presentó otra explicación con respecto a la formación de montañas. Sostenía que la Tierra era originariamente más achatada en los polos que ahora. Cuando se volvió más esférica, el abultamiento ecuatorial se contrajo. Esto originó presiones laterales que han generado montañas de plegamiento en dirección este-oeste. Darwin explicaba la existencia de cadenas norte-sur como efecto de mareas lunares sobre la rotación de la Tierra. Las fuerzas combinadas producirían pliegues en la corteza terrestre: pliegues que estarían situados en ángulo recto respecto de la dirección de mayor presión.

"En el caso de la Tierra", escribió, "los pliegues correrían de norte a sur en el ecuador... Una vez formado, cualquier pliegue tendría tendencia a desviarse levemente, situándose más de este a oeste que cuando se formó primitivamente."

Esta teoría tiene en la actualidad pocos sustentadores. Las tremendas fuerzas que Darwin menciona podrían haber actuado efectivamente sólo en la juventud del mundo, antes de la formación de las cordilleras. Además, la teoría de Darwin no explica la evidente relación entre sedimentos y montañas.

Millones de años de sedimentación precedieron el surgimiento del mar de los Alpes y los montes Cárpatos de Europa. La misma sucesión de sedimentación seguida de elevación es evidente en las cadenas de montañas del continente norteamericano. El geólogo estadounidense James Dwight Dana observó que "la región sobre la cual formaciones sedimentarias estaban desarrollándose para convertirse finalmente en los montes Apalaches llegaba desde Nueva York a Alabama, y tenía un ancho de 150 a 300 kilómetros... La pila de lechos horizontales medía en el centro 12.000 metros de profundidad. La pila de las montañas Wasatch tenía 18.000 metros de espesor... Los lechos de los Apalaches no se formaron en un océano, sino en aguas poco profundas donde estaba produciéndose un asentamiento gradual; y por último, cuando estaban dispuestos para la génesis (es decir, para el verdadero proceso de formación orogénica), descansaban en una depresión de 12.000 metros de profundidad, llenándola de borde a borde". La cadena Laramie, que se extiende desde el sudoeste de Wyoming hasta Colorado, es una estribación del sistema de las montañas Rocosas. Se estima que fue elevándose desde un sedimento a 15.000 metros de profundidad.

LA HIPÓTESIS DE BABBAGE

De modo que cualquier explicación de las cadenas de montañas debe tener en cuenta sus orígenes sedimentarios. Se han propuesto varias teorías para explicar cómo surgieron las montañas a partir de formaciones sedimentarias. Un geólogo inglés llamado Babbage ofreció una que en su momento tuvo considerable apoyo. Según Babbage, la acumulación del sedimento en el lecho oceánico debe causar una elevación de la temperatura del fondo primitivo, puesto que la temperatura aumenta con la profundidad. Un depósito de sedimento de 300 metros en el fondo del océano haría que la del área que cubría subiera hasta unos −7° C aproximadamente. Al mismo tiempo, la erosión de las superficies terrestres por el viento, el agua corriente, los agentes atmosféricos y otros factores causaría una disminución de la temperatura en otras áreas bajo la superficie de la Tierra, ya que no se hallarían a tanta profundidad como antes.

Babbage sostenía que el calor desigual de áreas adyacentes explicaba la formación de los sistemas de montañas. Las áreas cuya temperatura se había elevado como consecuencia de la sedimentación resultarían debilitadas y se verían más afectadas que antes por la presión lateral ejercida por las capas adyacentes cuyas temperaturas descenderían. Por ello, la presión empujaría hacia arriba las áreas calentadas.

El geólogo inglés T. Mellard Reade propuso una modificación a la teoría de Babbage. Sugirió que el sedimento fraguado sobre el fondo del mar actuaría como una faja de material aislante sobre la superficie de una bola de metal en enfriamiento. Retendría y elevaría el calor de la parte de la superficie de la Tierra que cubriera. Con el tiempo, esta faja sedimentaria se calentaría por encima de la temperatura promedio de la corteza. El exceso de calor produciría la expansión del sedimento y de la roca subyacente. La expansión lateral o hacia abajo sería restringida. De ahí que el área afectada se viera empujada hacia arriba "como se expande hacia arriba un pastel al hornearse".

¿ACTUÓ LA RADIACTIVIDAD?

Otro geólogo, John Joly, sostuvo que los materiales radiactivos acumulados en la corteza terrestre originaron una elevación de temperatura que la debilitó en muchos sitios y causó el plegamiento que ha producido la formación de montañas. "Dada una fuente local de calor [radiactividad] aplicada encima, mientras el calor normal de la Tierra fluye de abajo hacia arriba, el área donde estas condiciones existen debe necesariamente convertirse en el primer lugar que se ablande y ceda, tan naturalmente como se rompe una cadena por el eslabón más débil."

ISOSTASIA

En cierta época fue idea muy generalizada que la formación de montañas se debía a la *isostasia* (de dos palabras griegas que significan "igual estabilidad"). La isostasia es todavía un concepto válido en cuanto se refiere al mantenimiento del equilibrio en la corteza terrestre. Esto se logra cuando movimientos en gran escala en una zona están contrarrestados por otros compensatorios en otras zonas. Según el concepto de isostasia, los segmentos o bloques de la capa más externa de nuestro planeta están más o menos equilibrados entre sí y pesan aproximadamente lo mismo. Las tierras altas y las montañas están sustentadas por rocas de menor densidad que las situadas por debajo de las tierras bajas y del fondo del océano. Las montañas y llanuras puede decirse que "flotan" en una capa de roca aún más densa que tienen debajo. Esta capa profunda se ha comparado a un líquido, aunque en realidad es un sólido plástico que fluye o se deforma muy lentamente bajo altas presiones.

Las mesetas se asemejan a grandes icebergs, y las tierras bajas, a masas de hielo que flotan en agua más densa que ellas. Ciertamente, un témpano grande es en sí mismo mayor y más pesado que un témpano pequeño. El primero, por consiguiente, se hunde más y tiene una gran "raíz" de hielo que penetra en el agua

Georg Gerster. Photo Researchers

En muchas zonas montañosas ocurren fuertes movimientos horizontales de tierra. En tales áreas suelen producirse fracturas o fallas. Parte de la espectacular falla de San Andrés, en California, se ve a la derecha.

y la desplaza. El témpano grande, pues, está sostenido por una columna de agua mucho más corta que la que soporta el témpano chico. En la corteza terrestre, el ajuste isostático por flujo de roca profunda es una respuesta a fuerzas no isostáticas que causan desequilibrio y tensión en otras rocas.

La teoría más antigua de la formación de las montañas por la isostasia sostenía que el fondo de un mar se estaría llenando con sedimentos erosionados de una cadena de montañas adyacente. Al acumularse y presionar en el fondo, éste se curvaría hacia abajo formando una depresión que alcanzaría el sustrato cortical o *zona de flujo*. Las rocas del sustrato se desplazarían hacia las montañas para compensar la depresión en hundimiento y el aligeramiento de las montañas a causa de la erosión. Éstas se elevarían, pero siempre menos en cada sucesiva elevación. Habría también un límite en el descenso del geosinclinal y así disminuiría el flujo subcortical. El traslado de vastas cantidades de sedimento a través de la superficie de la Tierra era considerado uno de los principales factores de la formación de montañas. Este traslado, más las reacciones isostáticas provocadas por él, se suponía que habían producido las cordilleras.

Cuando los sedimentos, la materia prima de la orogénesis, se tornan uniformemente inclinados, el resultado es una formación llamada "lomo de cerdo", como se ve abajo.
Jerome Wyckoff

Hoy día se juzga inadecuada la isostasia, tanto para desarrollar geosinclinales como para producir orogenia. Por ejemplo, ciertas hoyas muy profundas en el fondo oceánico están sustentadas por rocas menos densas que las de tierras montañosas cercanas. Evidentemente, estas rocas que aquí no soportan todavía cargas pesadas de sedimento no han sido abatidas hasta ahora por movimientos isostáticos. Otras fuerzas, cuya naturaleza aún se discute, deben haber sido las que provocaron esta situación.

La isostasia puede, sin embargo, liberar fuerzas formadoras de montañas. A medida que las rocas del geosinclinal se sumergen en la zona de flujo, el calor de esa región las ablanda y las debilita. Las rocas más firmes que rodean las capas del geosinclinal a un nivel algo más alto pliegan y fracturan aquéllas por presión lateral. Por ello, muchas de estas capas se vuelven más gruesas. Por otra parte, las más bajas se sumergen aún más en la zona de flujo. Estos estratos más bajos forman una raíz de material liviano empujado fuertemente hacia abajo dentro de la parte más densa de la corteza terrestre. No obstante, existe un límite respecto de hasta dónde puede llegar. Por fin, la masa plegada puede tender a elevarse o "flotar" más arriba. También continuas presiones horizontales pueden al fin forzar los encogidos lechos a "montarse" o ser arqueados hacia arriba contra las rocas más resistentes que los rodean o sobre ellas. El material rocoso fundido que emerge a través de las capas sedimentarias probablemente contribuye también a elevarlas.

LA TEORÍA DE LA DERIVA CONTINENTAL

La teoría de la *deriva continental*, en un tiempo completamente desacreditada, ha ganado ahora amplia aceptación. Según la misma, vastas masas de tierra se han deslizado horizontalmente por largas distancias. Al hacerlo, han chocado con otras, produciendo la elevación de la corteza terrestre en el lugar de la colisión. La teoría sostiene, por ejemplo, que los Alpes y demás montañas del sur de Europa se formaron cuando el continente africano se corrió hacia el norte y tropezó con las masas de tierra allí situadas. Las montañas Rocosas en América del Norte y los Andes en América del Sur fueron el resultado del movimiento de ambas Américas hacia el oeste. Las cadenas montañosas del sur de Asia surgieron cuando el continente asiático derivó hacia el sur.

El cerro Aconcagua, la más alta montaña del hemisferio occidental. Los efectos de la erosión son evidentes en la apariencia aplanada y esculpida de la cresta.

Mario Fantin, Photo Researchers

Los sustentadores de la teoría de la deriva continental señalan que los contornos de las costas occidentales de los continentes del hemisferio oriental coinciden con los de las costas orientales de los continentes del hemisferio occidental. Esto sugiere que dichas regiones formaron en un tiempo una sola masa de tierra. Los estratos y fósiles de ciertos períodos geológicos en los ahora separados continentes muestran sorprendentes similitudes, lo cual parecería reforzar la idea de la deriva continental. Existen también pruebas de que en otras épocas los polos magnéticos de la Tierra se hallaban situados en posiciones muy distantes de las actuales.

Los viejos teóricos de la deriva continental sostenían que los continentes flotaban libremente en un sustrato universal de roca plástica o semilíquida. Luego, otros afirmaron que corrientes que brotaban del interior caliente de la Tierra movían las masas terrestres corticales de un lado para otro. Esto las desmembraba y plegaba sus estratos, provocando la formación de montañas.

Según los opositores de la teoría de la deriva continental, las pruebas proporcionadas por los estratos rocosos de la Tierra y los fósiles sugieren miles de millones de años de ordenado desarrollo continental. Esto hubiera sido imposible si las masas terrestres se hubieran movido de acá para allá. Además, ciertos fenómenos parecen indicar que los cimientos de los continentes penetran muy profundamente en la Tierra. Esto haría difícil de aceptar el concepto de su traslado, según los geólogos antiderivistas. Ellos creen que gran parte de las llamadas pruebas de la deriva continental puede explicarse de otro modo.

EROSIÓN DE LAS CORDILLERAS

Tan pronto como surgieron las nuevas montañas, los efectos de la erosión por el agua y el aire comenzaron a tallarlas en forma de picos, ahora ni la mitad de altos de lo que fueron originariamente. El viento, la lluvia, la escarcha, el hielo, la nieve y la poderosa fuerza del agua corriente han actuado incansablemente durante millones de años y el desgaste ha sido inmenso; entre 4.500 y 6.000 metros han sido eliminados de las viejas montañas de Gales. Más de 7.600 metros –poco menos de ocho kilómetros– han sido arrancados de las montañas del Distrito Lacustre de Inglaterra y de la cadena de los Apalaches en la parte oriental de los Estados Unidos. En un tiempo existió una capa sedimentaria de 15.000 a 18.000 metros de espesor sobre lo que ahora es el Simplón, un paso alpino que alcanza una altura aproximada de 2.000 metros. La cadena Laramie en las montañas Rocosas tendría hoy un promedio de 10.000 metros de altura si no fuera por la erosión sufrida en diversas proporciones.

Es interesante señalar que las crestas originales o anticlinales se desgastan mucho más de prisa que los barrancos, valles o sinclinales. La razón parece ser que los anticlinales están más expuestos a las fuerzas erosivas.

Los anticlinales son erosionados tanto más rápidamente que los sinclinales, que en algunos casos la relación original entre estos dos tipos de plegado se ha trastrocado. El sinclinal se

ha convertido ahora en la cresta de la montaña, y el anticlinal, en el valle. Esta inversión, por ejemplo, ha ocurrido en los montes Apalaches. Las cumbres de esta cordillera fueron primitivamente sinclinales, y los valles, anticlinales.

El grado de desgaste de las montañas y la forma que han tomado dependen de diversos factores. Uno de ellos es la dirección seguida por los estratos que las componen. Otro es el material de que están hechas. La piedra caliza se desgasta muy rápidamente y en general adopta una forma curva. En ciertos casos puede tomar el aspecto de "mampostería en ruinas que sugiere la idea de murallas derrumbadas y torres tambaleantes". Aunque el granito resiste la acción de los elementos más tiempo que la piedra caliza, también está sujeto a erosión y frecuentemente es desgastado en forma de enormes bloques cuadrados. Las montañas de granito rara vez tienen picos agudos. Habitualmente son redondas y macizas, como el pico Pikes, cuyo coronamiento es de granito puro. Las elevadas cumbres dentadas de los Alpes europeos, que constantemente desafían la habilidad y el valor de los alpinistas, están compuestas de duras rocas cristalinas de gneis y micasquisto. Estas altas y áridas elevaciones son de formación más reciente que las cadenas de bajas cimas redondeadas y pendientes uniformes, como las Adirondacks.

Los pináculos del monte Blanco, la montaña más alta de Francia, son ejemplos de formaciones de roca ígnea, despojadas y meteorizadas por millones de años.

F. Tairraz

Las crestas nevadas de las montañas Hindu Kush, en el Asia central, que en algunos sitios superan los 7.000 metros sobre el nivel del mar, son de origen más reciente que los Himalaya, de pendientes más suaves. Aunque un único pico de estos últimos, el monte Everest, alcanza casi 9.000 metros sobre el nivel del mar y es la cumbre más alta del mundo, la cadena del Himalaya es en su mayor parte considerablemente más baja que el Everest.

La altura promedio a lo largo del eje central de los picos del Himalaya no pasa de 6.000 metros sobre el nivel del mar. Esta cordillera es el resultado de una larguísima serie de levantamientos y depresiones, con una historia geológica que se remonta a los tiempos precámbricos en los estratos más viejos. En este antiquísimo proceso, la abrupta elevación del Everest es de origen comparativamente reciente.

Esto nos enseña que dentro de la historia de una particular cadena de montañas la intensidad de los movimientos ascendentes puede variar de modo considerable. Así, en los Himalaya existen tres períodos de desarrollo claramente definidos, comenzando en el norte y operando hacia el sur. Es probable que el proceso inicial fuera una violenta contracción de la corteza terrestre a lo largo del borde meridional de la gran meseta del Asia central. El carácter gradual de los cambios subsiguientes está indicado por el curso regular de los ríos Indo y Brahmaputra, que labran profundas gargantas a través de las montañas.

Los cambios que ocurran en cualquier área determinada pueden causar finalmente efectos generales sobre toda la cordillera. Así, el proceso de *denudación* –el despojo y desgaste de las cimas de las montañas por los elementos– puede estar compensado por una gradual elevación de la altura de la base causada por las fallas y los plegamientos de la corteza terrestre. Los geólogos que propugnan la teoría de la isostasia (que, como hemos visto anteriormente, subraya el equilibrio dinámico de fuerzas) sostienen que los efectos de la denudación se compensan con la elevación de la base. En

De acuerdo con la teoría de la deriva continental, estas montañas, los Andes, en Perú, se formaron al deslizarse las grandes masas terrestres.

Suecia, por ejemplo, gran parte de la tierra está surgiendo gradualmente del mar. En algunas zonas, los bosques de pinos de las laderas de las montañas han sido elevados más allá de los límites de la región boscosa hasta las regiones de nieves perpetuas, donde el frío ha matado los árboles. Las costas de Escocia son notables por sus terrazas de roca, que se alzan una sobre la otra hasta una altura superior a 20 metros sobre el mar.

Este incremento de la elevación de la base es frecuentemente acompañado por la intrusión, en la parte inferior de la montaña, de grandes *batolitos*, rocas ígneas con un núcleo resistente de granito. Al contrario que las rocas efusivas, las masas intrusivas de roca ígnea rara vez llegan a la superficie de la Tierra, pero pueden actuar como pesada palanca para elevar los estratos de la superficie. Es posible, sin embargo, que una masa intrusiva situada cerca de ésta pueda quedar con el tiempo al descubierto por la acción de los elementos.

EL CALOR INTERNO EN FUNCIONES

La roca ígnea es, como su nombre indica, roca fundida del interior de la Tierra. El calor interior del planeta está constantemente ejerciendo una presión contra los estratos superiores de roca, creando pliegues y fallas en la corteza, a través de los cuales presiona la roca ígnea. La roca fundida original se enfría y solidifica en el proceso de alcanzar las capas superiores. Se distribuye en varias formas, tales como los filones auríferos a dos aguas del sur de Australia o las bóvedas y cúpulas que se encuentran en las regiones de estaño del oeste de Cornualles.

EDAD DE LAS MONTAÑAS

Al estimar la edad relativa de las montañas y los períodos variables de su formación, los geólogos emplean ciertos patrones. Los fósiles marinos ya han sido mencionados, y éstos son útiles dentro de límites definidos. Resulta evidente, sin embargo, que la Tierra es mucho más antigua que los fósiles marinos disponibles. Éstos son formas complejísimas, resultado de un largo proceso evolutivo, y no se hallaban presentes en los comienzos de nuestro planeta.

Para suplementar los fósiles se apela a la prueba de los elementos radiactivos en las rocas. Todos los minerales radiactivos que se hallan en éstas son una forma de reloj geológico. Existe una proporción constante mensurable de desintegración del elemento radiactivo que determina, dentro de estrechos límites, la edad de las rocas. Este maravilloso reloj autooperante revela que la Tierra tiene varios miles de millones de años, muchos más de los que antaño se había supuesto.

LLANURAS, MESETAS Y DESIERTOS

Entre las regiones más notables del mundo se encuentran las llanuras, las mesetas y los desiertos, vastas extensiones de tierra sin accidentes de consideración. Entre ellas hay desde planicies situadas por debajo del nivel del mar, hasta mesetas que se encuentran a 4.500 metros o más sobre éste. No existe una distinción muy marcada entre llanuras y mesetas. Por lo general, las primeras son regiones relativamente planas situadas a menos de 300 metros de altitud, y las segundas, extensiones similares de tierra elevadas más de 300 metros sobre el nivel del mar. Empero, lo que determina que una extensión plana de tierra sea una llanura o una meseta no es, necesariamente, su altura absoluta, sino su elevación comparada con la del área circundante.

La palabra desierto se aplica de manera vaga e indefinida a las llanuras y mesetas áridas y semiáridas. Muchos de los llamados desiertos no lo son en realidad, ya que, anualmente, en determinadas estaciones, reciben unos pocos centímetros de lluvia, que les permiten sustentar algo de vegetación. El desierto de Kalahari, en África, es un ejemplo: es un semidesierto. Muchos verdaderos desiertos, como el Sahara, en África, y el de Arabia, en Asia, pueden no recibir ni una gota de lluvia y carecen totalmente de vegetación.

VARIEDAD DE CLIMAS

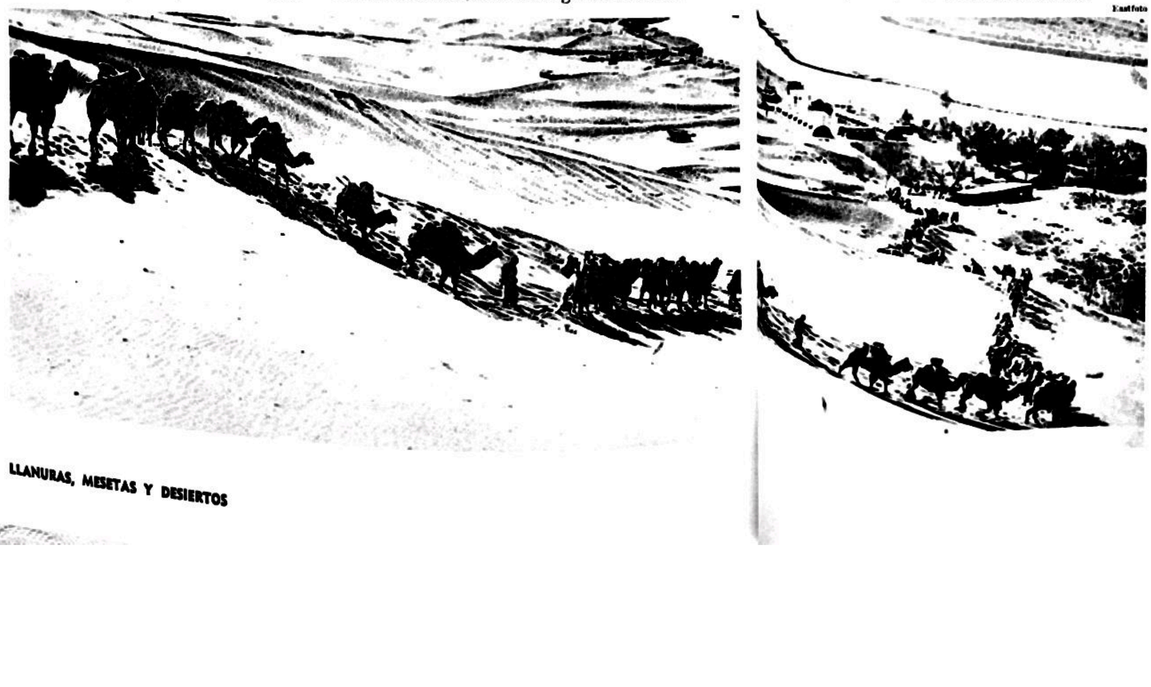
Las condiciones climáticas de las llanuras y mesetas son más bien variadas. De acuerdo con su latitud, situación y elevación, varían de calientes a frías y de secas a húmedas. Muchas son frías y secas, cualquiera sea su situación, a causa de su altitud, como sucede con la meseta del Tibet y el Altiplano de Bolivia. La aridez de algunas llanuras y mesetas se debe a una serie de factores, como la proximidad de altas cadenas de montañas, el calor y las corrientes oceánicas adversas. La circulación atmosférica también influye. Los Andes, que separan a Chile y la Argentina, interceptan en una parte los vientos húmedos del Pacífico. Al chocar contra la cordillera, éstos descargan su humedad en forma de lluvias que fertilizan las estribaciones chilenas, mientras las argentinas permanecen áridas. Lo contrario ocurre más al norte: los vientos alisios que cruzan la cuenca del Amazonas permiten que crezcan selvas tropicales en las laderas orientales de los Andes, pero en las occidentales empieza una de las regiones costeras más secas del mundo. Entre las extensas praderas productivas se cuentan las pampas argentinas, las estepas de la Unión Soviética y las Grandes Planicies de los Estados Unidos. Son regiones agrícolas y ganaderas de gran importancia.

ORÍGENES

Las llanuras y mesetas son el resultado de variados procesos geológicos. Un factor muy importante es la erosión provocada por el agua y el viento, y el desgaste de ciertos terrenos, incluso de altas montañas, hasta producir extensiones llanas o casi llanas. Esas planicies, o semiplanicies, llamadas *penillanuras*, pueden ser elevadas por las fuerzas terrestres hasta formar una meseta. Esta última sufre luego los efectos de la erosión, que la va rebajando progresivamente, a menos que sea levantada una vez más. Si las rocas que componen la meseta

El desierto de Gobi es el más extenso de Asia. Aquí una caravana de camellos se aproxima a uno de los pocos lagos pequeños que se encuentran en esa zona.

Eastfoto



Sovfoto

Las tundras son llanuras desoladas, bajas y frías. Algunos alces corren aquí sobre la tundra estéril y cubierta de nieve en el norte de la Unión Soviética.

varían en dureza y posición, son labradas hasta formar colinas y montañas. Ciertas formaciones montañosas del mundo actual son los restos de las que una vez fueron mesetas muy extensas. Esto ha sucedido, por ejemplo, con las montañas Catskill del estado de Nueva York.

Las llanuras costeras –tierras bajas que bordean las costas de algunos continentes, especialmente Europa y América del Norte– son, por lo general, sectores adyacentes del suelo submarino que han sido levantados en épocas geológicas recientes y forman ahora parte de la tierra firme. Su planicie refleja los lechos normalmente lisos de sedimentos que las componen y el efecto erosivo de las olas del océano. Las llamadas *llanuras aluviales* están formadas por sedimentos depositados por los ríos sobre la tierra. Las *planicies de los deltas* como la del Nilo, se extienden sobre los vastos depósitos que los grandes ríos dejan en sus desembocaduras. Un valle fluvial excepcionalmente ancho y llano es considerado muchas veces como una llanura. Las *planicies lacustres* se forman sobre lechos de lodo o de piedra que han dejado lagos hace mucho tiempo evaporados. Vemos así cómo una serie de agentes geológicos y climáticos pueden contribuir a formar llanuras, mesetas y desiertos.

La extensión de las tierras llanas de la corteza terrestre puede oscilar entre unos pocos y varios millones de kilómetros cuadrados. Cuanto más extensas sean estas llanuras, tanto más

El nombre "piedmont" se ha adoptado para designar una meseta del tipo de la meseta Piedmont, cuyas características la ubican entre una zona baja y una montañosa. En la ilustración de abajo se ve un típico perfil piedmont. La foto de la derecha muestra las ricas tierras de cultivo de la zona.

Hammond Inc.

Max Thorpe · Foto Alpha Asociados

Hammond Inc.

Oficina Nacional de Turismo Español

La meseta central de España constituye un ejemplo típico de las mesetas que suben más o menos abruptamente desde el mar o desde una estrecha llanura costera. Arriba, a la izquierda, vemos un perfil de ese tipo de meseta. Arriba, a la derecha, un paisaje típico de la meseta española.

fácil es que se aparten de la idea corriente de una superficie prevalentemente plana o de nivel parejo. Así, por ejemplo, aunque el extenso Sahara sea por lo general llano u ondulado, en su mismo centro tiene elevadas cadenas de montañas y dunas arenosas.

LLANURAS

La llanura más extensa del mundo es la de Eurasia, que se extiende a través de la región septentrional de Europa y de Asia, desde el mar del Norte hasta el este de Siberia. Los montes Urales dividen esta inmensa extensión en dos partes, una oriental y otra occidental. Otra gran llanura es la cuenca amazónica, que presenta el declive más suave del mundo.

Otra se extiende hasta el centro de América del Sur, bordeando la base de los Andes, e incluye una quinta parte del continente.

Landas. Algunas llanuras europeas reciben el nombre de landas. Las landas del sudoeste de Francia son llanuras pantanosas situadas en los departamentos de Landes, La Gironde y Lot-et-Garonne. Son de origen submarino y tan cenagosas que los pastores de la zona caminan sobre zancos para reunir sus rebaños. Las landas francesas cubrían originalmente cerca de 800.000 hectáreas, gran parte de las cuales han sido mejoradas.

Aún más extensas son las landas de Holanda y del norte de Alemania. En Holanda hay 1.000.000, o más, de hectáreas cubiertas de esponjosos musgos de pantano.

Estepas. En Hungría y la parte central de la

Hammond Inc.

Eastfoto

La meseta del Tibet es un ejemplo del tercer tipo de meseta, aquella que está rodeada de montañas. Arriba, a la izquierda, el perfil de una meseta de ese tipo. Arriba, a la derecha, una foto de la llanura de Lhasa, que forma parte de la meseta del Tibet.

Unión Soviética, en lugar de landas hay llanuras conocidas con el nombre de estepas. En Hungría son simplemente praderas cubiertas con pasto natural y flores silvestres sobre las que pacen manadas de bueyes y caballos. En el centro de la Unión Soviética las estepas son enormes; las de Chernozom ocupan alrededor de 80.000.000 de hectáreas, cubiertas en su mayor parte de pasto natural; otra gran extensión está sembrada de maíz. La tierra vegetal es rica y difícil de superar para el cultivo de este cereal. Al sur de las estepas de Chernozom, en las cercanías del mar Caspio, hay otras estériles de arena blanca y arcilla rojiza, donde sólo crecen unos pocos arbustos, y hasta una distancia de unos 500 kilómetros en línea recta sólo se encuentra un solo tipo de árbol, que es una especie de álamo.

Tundras. Las llanuras bajas del norte de la Unión Soviética se denominan tundras. Constituyen una de las regiones más desoladas del mundo. En invierno están nevadas, y en verano se convierten en tierras pantanosas, cubiertas de una espesa capa de musgo rengífero y líquenes blanquecinos.

Praderas. Hay una gran extensión de terreno en el Valle del Misisipí casi desprovista de árboles, denominada ***región de pradera.*** Su límite oriental está constituido por una línea boscosa irregular que corre en dirección sudeste, a través de Minnesota y Wisconsin hasta el oeste de Indiana, volviéndose luego hacia el sudoeste a través de Illinois y Misurí hasta los límites de Oklahoma y la parte oriental de Texas. Desde esta línea, las praderas se extienden de 300 a 950 kilómetros hacia el oeste, hasta unirse a las Grandes Llanuras. Se cuentan hoy entre las regiones agropecuarias más importantes de los Estados Unidos.

Pampas y llanos. En América del Sur, en las cuencas del Amazonas y otros grandes ríos hay llanuras cubiertas alternativamente de selvas y tierras de pastoreo similares a las de la cuenca del Misisipí. Hay también zonas totalmente cubiertas de hierba. En América del Sur, las llanuras herbosas reciben el nombre de pampas y llanos. En las zonas más cálidas crecen entre el pasto mimosas y otros arbustos. Los llanos de Venezuela y de Colombia se caracterizan porque son alternativamente desiertos o tierras de pastoreo. Antes de la estación de las lluvias el suelo se seca, toda la vegetación perece, y los llanos se convierten en un verdadero desierto. Luego, repentinamente, las tormentas de la estación lluviosa inundan el suelo. Una enorme cantidad de plantas surge del polvo, y la gran extensión amarilla se trasforma en una pradera florida. Las llanuras más notables de América del Sur son las pampas argentinas, extendidas sobre más de 1.500 kilómetros entre el Brasil y la Patagonia. Allí se encuentran, probablemente, los campos de pastoreo más importantes del mundo.

MESETAS

Las mesetas pueden dividirse en tres clases diferentes. Algunas, como la meseta Piedmont de América del Norte, separan una llanura baja de una zona montañosa. Otras, como las del Tibet, están rodeadas de montañas. Y finalmente otras, como la meseta central de España, se elevan más o menos abruptamente desde el mar o desde estrechas llanuras costeras.

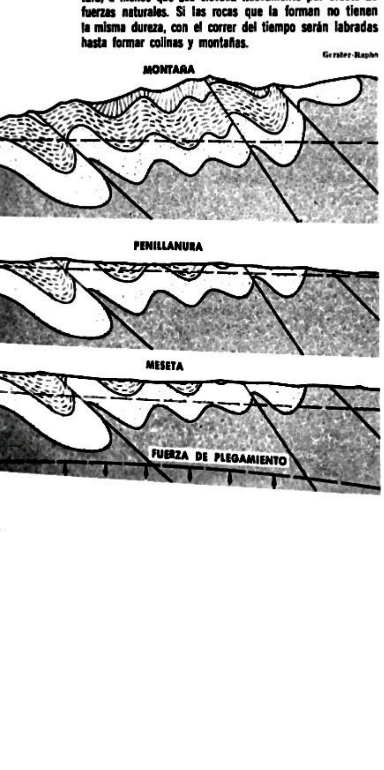
Los Estados Unidos tienen varias mesetas

que pertenecen a la primera y segunda de las clases mencionadas. La Piedmont está al oeste de la llanura costera atlántica y al este de los montes Apalaches. Se extiende desde Nueva Jersey hasta Alabama, alcanzando un ancho de casi 500 kilómetros en Carolina del Norte. La meseta de Colorado se extiende a través de 330.000 kilómetros cuadrados sobre zonas de Arizona, Utah, Colorado y Nuevo México. Está cortada por varios cañones muy profundos, incluyendo el célebre Gran Cañón del Colorado. En el noroeste del país se encuentra la meseta de Columbia, que abarca Oregón, Washington e Idaho, en una extensión de unos 250.000 kilómetros cuadrados. Gran parte está formada por sucesivas coladas de lava.

Algunas mesetas, como las de Pamir y del Tibet, se encuentran entre las regiones más inaccesibles del mundo. Otras, como la de Uganda, tienen un clima estimulante, que ofrece muchas veces contraste muy marcado con el de las regiones vecinas. En las Américas, las civilizaciones de los aztecas, toltecas e incas estaban concentradas sobre altas mesetas, y en la actualidad todavía se encuentran allí muchas ciudades importantes.

Tres pasos en la formación de un tipo de meseta. Arriba: las montañas se forman a consecuencia de las fallas y plegamientos. Centro: la erosión desgasta gradualmente las montañas hasta convertirlas en una penillanura. Abajo: los plegamientos hacen que la penillanura se eleve hasta formar una meseta. La meseta a su vez será atacada por la erosión y disminuirá gradualmente su altura, a menos que sea elevada nuevamente por efecto de fuerzas naturales. Si las rocas que la forman no tienen la misma dureza, con el correr del tiempo serán labradas hasta formar colinas y montañas.

Gerster-Rapho



Las vastas mesetas en la vecindad del sistema montañoso del Himalaya, en Asia, sirven para "prolongar el norte hasta el corazón mismo del sur y para unir, dentro de un espacio limitado, todos los climas del año". La meseta más alta es la del Tibet, con una altitud media de 4.500 metros. Está rodeada al sur por los poderosos Himalaya, y al norte por las montañas Kunlun. Otra gran meseta es la de Pamir, a la que los nativos llaman el "Techo del mundo". Se halla a una altura de unos 3.500 a 4.000 metros sobre el nivel del mar. La meseta del Irán cubre unos 2.500 kilómetros cuadrados sobre partes de Irán, Afganistán y Beluchistán; su altura media oscila entre los 90 y los 1.500 metros.

Las mesetas de Europa no pueden compararse en altitud con las del Tibet o de Pamir. Más de la mitad de España está formada por una cuya altura media es de 600 metros. Gran parte de Baviera, en Alemania, es también una meseta cuya altura media es de alrededor de 500 metros sobre el nivel del mar.

Gran parte del continente africano, cuya extensión es sólo inferior a la del asiático, está formado por mesetas cuya altitud media es de 600 metros. Una sucesión de altas mesetas corre desde el mar Rojo hacia el sur y hacia el oeste por miles de kilómetros. Aunque la mayor parte del continente está compuesta por ellas, tiene también altas cadenas de montañas, el vasto Sahara, el bajo valle del Nilo y las regiones costeras.

Una meseta sudamericana muy interesante es el Altiplano, ubicado principalmente en Bolivia, entre dos ramales de la cordillera de los Andes. El Altiplano, que está alrededor de 3.650 metros de altitud, tiene 725 kilómetros de largo y 130 de ancho; es una de las mesetas altas más pobladas del mundo. Las montañas que la rodean la encierran de tal manera que toda el agua de lluvia queda retenida en su propia cuenca y drena hacia sus propios lagos.

DESIERTOS

Además de las landas, tundras, praderas, llanos y estepas, están los grandes desiertos propiamente dichos, que algunas veces son llanuras y otras mesetas. El desierto conocido, y quizá el más típico, es el Sahara, que se extiende a través de África, desde el Atlántico al mar Rojo, con un largo de alrededor de 5.000 kilómetros. El ancho promedio de esta zona desértica supera los 1.700 kilómetros, y en total cubre una extensión tan grande como todo el continente europeo. Su altura media es de 450 metros y, por tanto, puede considerarse más una meseta que una llanura, salvo por el hecho de que no hay tierras bajas que contrasten con el nivel general de las arenas. Por extraño que parezca, sólo alrededor de una octava parte de la superficie del Sahara está compuesta de arena: la mayor parte es tierra y piedra.

…A QUE MUESTRA LOS PRINCIPALES DESIERTOS DEL MUNDO

Un grupo de altas montañas que se extiende de norte a sur, a lo largo de la zona media del Sahara, lo divide en una parte oriental y otra occidental. La parte occidental es un yermo de arena barrido por los vientos. En el oeste las arenas son llevadas al océano; en el sur, avanzan hacia los ríos Níger y Senegal, en cantidades tales que llegan a modificar sus cursos. El Sahara oriental es también una extensión de arena, pero muchas mesetas de rocas y de arcilla y buena cantidad de montañas rompen la monotonía de las llanuras.

El Sahara es tan caluroso y seco que en su mayor parte resulta casi inhabitable. La falta de agua es fatal para la mayoría de las formas de vida. Los únicos seres que abundan son escorpiones, lagartijas, víboras y hormigas.

Este desierto está sometido a temperaturas extremas. Se han registrado marcas de 58° C. Empero, en la estación fría el termómetro puede descender hasta unos 4° C. Suele producirse escarcha, particularmente en las regiones occidentales y en las tierras altas de las zonas centrales. La humedad relativa puede descender, en verano, hasta el nivel, extraordinariamente bajo, del cinco por ciento, aún en la proximidad de los oasis. En la estación fría puede llegar hasta el 47 por ciento. La temperatura suele descender treinta grados, o más, durante la noche hasta en la estación calurosa. El aire es tan seco que el calor de las arenas calcinadas por el sol durante el día se irradia rápidamente.

En cualquier lugar donde surge agua de la tierra, o llega al desierto desde las colinas, se forma un oasis –pequeña isla de verdor–, que resalta en marcado contraste con las deslumbrantes arenas del desierto circundante. Los oasis son muy numerosos en el Sahara. Están situados de manera tal que en algunas regiones forman una línea a través del desierto, lo que permite recorrer, por etapas, largas distancias. Allí el terreno es extraordinariamente fértil: sólo se necesita agua para hacerlo producir. Lamentablemente, no puede confiarse en la provisión de agua de los oasis, ya que a veces la sequía es muy marcada. En ocasiones, las corrientes de agua que llegan al desierto provenientes de las colinas se convierten en torrentes furiosos, que pueden arrastrar árboles y plantaciones de cereales.

La planta más común en los oasis es, por supuesto, la palma datilera, cuyo fruto, el dátil, constituye un alimento muy concentrado. A veces se comen tal como se arrancan de la planta; otras, se los conserva molidos y pren-

Gerster Rapho

sados. En los oasis del Sahara abundan damascos, duraznos, granadas, naranjas y otros tipos de frutas. Algunos cereales, como maíz, trigo y cebada pueden cultivarse bien en ellos, y el trébol florece cuando hay irrigación.

En el sur de Túnez y norte de Argelia, la mayor parte del Sahara está por debajo del nivel del mar, y en un tiempo fue lecho de una bahía llamada del Tritón por los antiguos escritores griegos. Una erupción volcánica separó esta bahía del Mediterráneo y, a consecuencia del intenso calor, sus aguas se secaron, dejando unos pocos lagos salados dispersos. El terreno de esta zona es todavía potencialmente muy fértil. Si la bahía del Tritón pudiera reconstituirse, permitiría en el Sahara una producción agrícola intensiva.

En este desierto se ensayaron distintos proyectos de irrigación. En muchos lugares, especialmente al sur de los montes Atlas y en Argelia, se obtuvo agua de pozos artesianos. En la actualidad se sabe que el Sahara tiene reservas inmensas de agua subterránea, suficientes para convertir zonas muy extensas en tierras fértiles. Gracias a ellas podría recuperar las condiciones de humedad de que disfrutó hace varios milenios.

El Sahara es un vasto desierto, con inmensas llanuras interrumpidas por mesetas y montañas. Izquierda: un oasis en la región del Sahara argelino. Surge entre las arenas como una pequeña isla de verdor que rodea un depósito de agua. Abajo: una tormenta de arena en el Sahara, que está sometido a condiciones climáticas extremas.

Barge

DESIERTOS DE ASIA

Atravesando el mar Rojo desde el Sahara se encuentran los desiertos de Arabia. El Nefud, situado en el norte, es un gran cinturón de dunas arenosas, cuyo ancho varía de 80 a 225 kilómetros. Estas dunas son a menudo muy alargadas y están separadas por valles que llegan a tener hasta 16 kilómetros de ancho. En la parte centrooriental de la península arábiga se extiende la desolada región conocida con el nombre de Dahana. Consiste principalmente en llanuras de grava, cubiertas a intervalos de arena roja. En esta región, los picos de arena llegan a veces a una altura de 90 metros. El desierto arenoso meridional de Arabia, llamado Rub' al Khali, es una región muy extensa, de 1.200 kilómetros de largo por 640 de ancho, donde depresiones de arenas duras alternan con zonas de otras blandas y movedizas.

El desierto de Gobi es el más grande de Asia. Su nombre chino es Han Hai ("mar seco"), o Shamo ("desierto de arena"). El nombre "Gobi" se aplica algunas veces a toda la extensión de tierra árida que va desde el Pamir hasta Manchuria y desde los montes Altai y Sayan hasta el Nan Shan, o sea, las montañas del sur de China. El desierto de Gobi propiamente dicho se encuentra en la parte oriental de esta inmensa región. Es una meseta desértica de grava y arena y su altitud oscila entre 900 y 1.500 metros. Tiene alrededor de 1.600 kilómetros de largo por 800 de ancho.

Los árboles son casi desconocidos en esta zona. La vegetación está compuesta principalmente por hierbas y arbustos, y el agua se encuentra sólo en pozos y en lagos pequeños, en cuyas márgenes hay algunos oasis. Las temperaturas extremas que se registran en el desierto de Gobi son tan impresionantes como las del Sahara. Durante el invierno, cuando los vientos fríos del norte barren la tierra, la temperatura cae algunas veces hasta 50° C bajo cero, y durante los meses de verano puede alcanzar hasta 45° C a la sombra.

DESIERTOS AMERICANOS

Posiblemente, la zona desértica más notable de América del Norte sea el Gran Desierto del Lago Salado, situado en el noroeste de Utah, región inhabitable del Gran Desierto Americano. Se extiende unos 180 kilómetros al sur

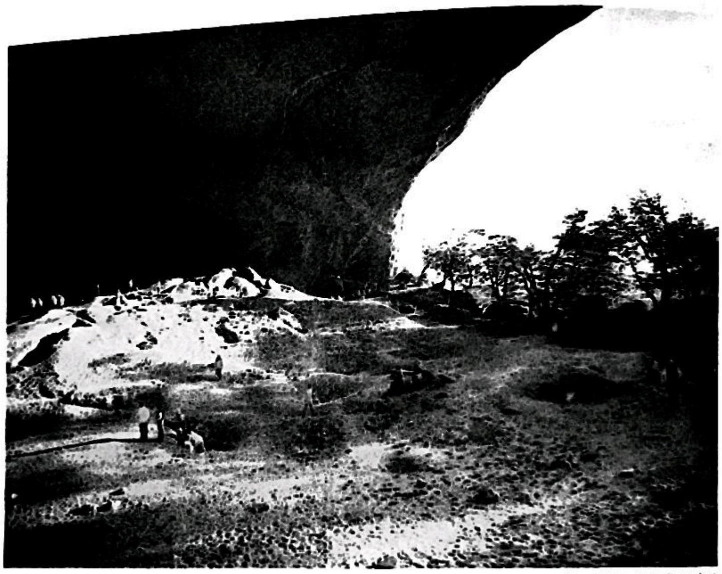
Fotografía del terreno seco, fracturado y calcinado de una típica región desértica.

de las montañas Grouse Creek, y está formado principalmente por vastas planicies de un gris blancuzco de las que emergen, como islas, algunos picos y cadenas aisladas. El suelo de esta planicie está sembrado de millones de toneladas de cristales de sal.

En América del Sur encontramos el desierto de Atacama, famoso por sus depósitos de nitrato. Esta región árida, situada en el norte de Chile, limita al oeste con la cadena costera del Pacífico, y al este con los Andes. Se extiende unos 950 kilómetros al sur desde la frontera peruana. En esa región prácticamente nunca llueve y es casi totalmente estéril. Pero en el siglo XIX los chilenos comprendieron el inmenso valor de sus casi inagotables yacimientos de nitratos. El desierto de Atacama fue la fuente más importante de este valioso elemento, hasta que comenzaron a elaborarse los nitratos sintéticos.

DESIERTOS AUSTRALIANOS

La mayor parte de la región interior de Australia, especialmente hacia el oeste, está formada por tierra árida. En el norte se encuentra el Gran Desierto Arenoso y en el sur, el desierto Victoria. En algunos lugares se cría ganado. Otras zonas, como las formadas por dunas fijas, nunca fueron habitadas. Allí las hierbas, arbustos y matorrales fijaron los médanos de arena.



Russ Kinne Photo Researchers

Las cavernas han tenido gran significación en el trascurso de la historia humana: marcas en sus paredes, restos de herramientas sencillas y otros indicios revelan que las familias prehistóricas las utilizaban como hogar; tal es el caso de esta grande, en Chile.

CAVERNAS

Las cavernas siempre han ejercido gran fascinación sobre la humanidad, principalmente, quizá, por los misterios que con tanta frecuencia las rodean. A medida que un explorador se arrastra a través de un estrecho pasaje, perforando la oscuridad con el haz de su lámpara, se pregunta qué hallará a la vuelta del próximo recodo. Quizás un río subterráneo, o un estanque calmo, o un foso muy abierto. Puede toparse con una antecámara con ondulantes cortinados de piedra, o una sala magnífica ornada con esculturas naturales que desafían la mejor creación del hombre.

En algunas cavernas se puede perder todo contacto con la realidad. Que alguien penetre, solo si es posible, en una cámara remota de una de estas cavernas, extinga las luces, se siente y aguarde. Nada hay que actúe sobre sus sentidos. Está en oscuridad absoluta. Ni un sonido, ni el más leve movimiento de aire. A medida que trascurre el tiempo, el cuerpo ya no le parecerá parte de sí mismo.

En otras cavernas, el silencio es quebrado por el sonido del agua que gotea, o por un arroyo borboteante. Puede haber una corriente continua de aire, como en la famosa Caverna del Viento de Dakota del Sur, en los Estados Unidos. Se puede oír el sonido sibilante de murciélagos que han hecho de la caverna su hogar, o ver, a la distancia, un pequeño rayo de luz procedente de la entrada.

FORMACIONES CALIZAS

La mayoría de las cavernas se encuentra en formaciones calizas de muchas diferentes partes del mundo. La caliza es una roca sedimentaria, fundamentalmente constituida por carbonato de calcio ($CaCO_3$), con diversas impurezas, esto es, sustancias tales como cuarzo, pedernal u óxido de hierro. También puede haber una proporción considerable de arcilla. La mayor parte de la caliza es resultado de la actividad orgánica; representa los restos de organismos marinos, como corales, moluscos y diversos animales y plantas microscópicos.

En todos los continentes se hallan formaciones calizas, cuyo espesor va desde unos pocos centímetros hasta muchos metros. En los Estados Unidos aparecen, principalmente, en un cinturón, que va desde Nueva York hasta Alabama, en los estados centrales de Ohio, Indiana, Kentucky y Tennessee, y en varios estados del oeste. Existen importantes formaciones de caliza en varias zonas de Canadá, incluyendo la región de Grenville, en el sur de Ontario; la de los Apalaches, en las tierras bajas del San Lorenzo, y en la del río Mackenzie.

El carbonato de calcio –del que está compuesta principalmente la caliza– cristaliza en dos formas: calcita y aragonita, que tienen diferente estructura interna. La calcita es mucho más común que la aragonita.

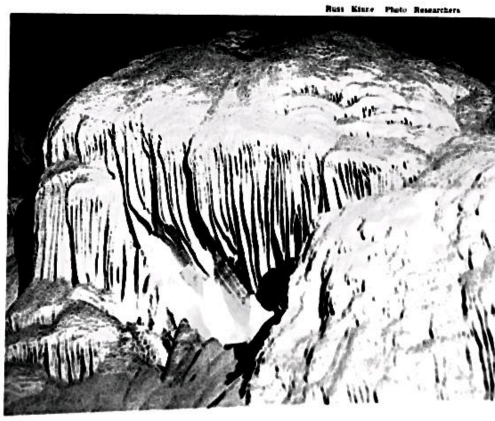
El carbonato de calcio sólo es soluble en agua que contenga disuelto el gas bióxido de carbono (CO_2), que se halla prácticamente en todas las aguas de superficie. Proviene de la atmósfera. La reacción química que se produce cuando el agua con bióxido de carbono actúa sobre el carbonato de calcio es la siguiente:

$$\underset{\text{agua}}{H_2O} + \underset{\text{bióxido de carbono}}{CO_2} + \underset{\text{carbonato de calcio}}{CaCO_3} \underset{\text{da}}{\rightarrow} \underset{\text{bicarbonato de calcio}}{Ca(HCO_3)_2}$$

El bicarbonato de calcio que resulta de la reacción es muy soluble y, en consecuencia, fácilmente eliminado por el agua, lo que explica por qué la caliza es una de las rocas más débiles en regiones húmedas. En las zonas secas, tales como el Gran Cañón, en E.U., es una de las formaciones más resistentes.

En los lugares de fuerte precipitación pluvial se disuelve una gran cantidad de caliza. Por ejemplo: la lluvia que cae en un solo año en una hectárea del suelo de la región de la Caverna del Mamut puede disolver uno y medio metros cúbicos, o más, de caliza.

Russ Kinne Photo Researchers



La acción del agua sobre la caliza puede crear algunos efectos muy llamativos, tales como los de la Boca de la Ballena, en las Cavernas Carlsbad, Nuevo México.

Ésta aparece en capas más o menos horizontales en la región de la Caverna del Mamut, en Kentucky, y en otras zonas de éste y otros estados vecinos. Esto sucede porque el antiguo lecho marino ha sido elevado en conjunto hasta su posición actual, a más de 100 metros sobre su nivel anterior. En otros lugares, tales como Dalmacia, en Yugoslavia, y en la región de las cavernas de Carlsbad, en Nuevo México, la actividad orogénica ha flexionado y encorvado la caliza en forma de pliegues. Tales lechos están inclinados en diversos ángulos y, con frecuencia, quebrados por fallas.

Veamos qué ocurre tanto encima como debajo del suelo cuando se forman cavernas en una región donde los estratos calizos son horizontales, como en la zona de la Caverna del Mamut. La Figura 1 muestra parte de la región de meseta cerca de la Caverna. El territorio es casi llano; los muy pocos ríos fluyen en valles de escasa profundidad. La superficie de la región está formada por una capa delgada de arenisca.

Las estalactitas y las estalagmitas son rasgos prominentes de las cavernas calcáreas. Las primeras parecen gotear desde el techo, mientras que las segundas crecen desde el piso. Cuando se juntan, pueden formar una columna.

Donde ésta ha sido desgastada aparece un suelo rojizo y amarillento, el residuo que quedó una vez disuelta la caliza formado por diversas impurezas acumuladas originalmente con ésta cuando se depositó sobre el fondo del mar. Debajo de la cubierta superficial hay una formación sólida de caliza bastante pura, representada en el grabado por una cuadrícula, que es el símbolo corriente de esta sustancia.

En la superficie del terreno hay varias hondonadas poco profundas más o menos circulares, llamadas sumideros, que se formaron por la acción erosiva del agua en las zonas donde hay grietas, o junturas, en las rocas. El agua superficial se abre camino desde estos sumideros hacia el suelo, y posiblemente hacia adentro de las grietas y de los *planos de estratificación* de la caliza. Un plano de estratificación es la superficie que separa una capa de roca estratificada de otra que tiene por arriba o por debajo.

El agua subterránea disuelve rápidamente canales a lo largo de las junturas verticales, o grietas, así como a lo largo de los planos de estratificación más solubles, y se forman las cavernas. Posteriormente, el agua escapa en manantiales por los flancos de los arroyos subterráneos.

Existen, literalmente, decenas de miles de sumideros en una zona como la de la Caverna del Mamut. Gradualmente se ensancha a medida que sus bordes se desgastan. Con el tiempo, se unirán varios para formar una depresión en forma de valle, o *sumidero de valle*. De esta manera, la región se descompone en sumideros y sumideros de valle, así como en verdaderos valles que contienen arroyos superficiales. Por debajo de la zona de meseta que está entre estos arroyos, la región se encuentra cribada con cavernas que tal como ilustra claramente la Figura 2, poseen diferentes niveles interconectados mediante pasajes estrechos. Aquí, el agua subterránea ha abierto, por disolución, canales tortuosos a lo largo de los planos de unión.

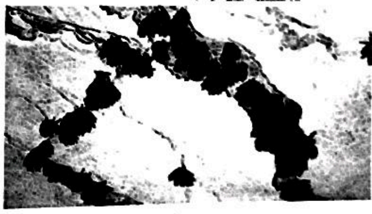
En el nivel más bajo de una caverna como la del Mamut puede haber un río subterráneo, que a veces surge del suelo como poderoso manantial y se une a uno de los ríos superficiales de la región. Por debajo del nivel de este arroyo de superficie, las grietas y pasajes en la caliza están completamente inundados. A esta profundidad hay poca actividad, porque el agua apenas se mueve.

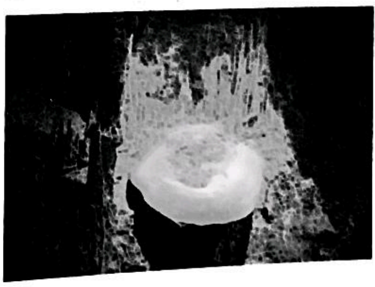
Sin embargo, por encima de este nivel, el agua de lluvia que se filtra desde arriba agranda constantemente las cavernas. En ciertos lugares, se evapora y deja en las cuevas depósitos cuyas formas fantásticas deleitan al visitante.

En regiones bajas, tales como Florida, Yucatán y Barbados, en las Indias Occidentales, el agua subterránea aparece cerca de la superficie. Muchos sumideros llegan a este nivel. Según las dimensiones de los mismos se forman manantiales, estanques o lagos. Por encima del nivel del agua subterránea hay innumerables cavernas y pasadizos pequeños.

No todas las grutas calizas son tan simples. Las famosas Cavernas de Carlsbad, por ejemplo, se han abierto en una formación inclinada y maciza de 300 metros de espesor, con un sus-

Con frecuencia, los murciélagos habitan las cavernas. Aquí, un grupo pende del techo de una mientras hibernan.





La interacción del agua y la caliza frecuentemente produce algunos efectos insólitos; esta "bóveda" es una formación que se encuentra en las Cavernas Mitchell, en el Desierto del Mojave, California.

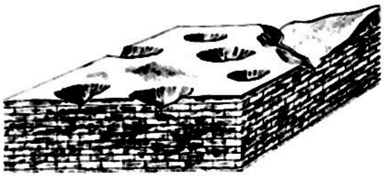


Fig. 1. Sección representativa de la región de meseta cercana a la Caverna del Mamut. El territorio casi plano sólo está atravesado por algunos ríos que corren por valles poco profundos. Existe una formación calcárea maciza (indicada en la ilustración por el cuadriculado) por debajo de la cobertura superficial. Los hoyos planos son sumideros que se formaron por la acción erosiva del agua.

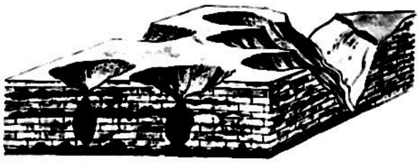


Fig. 2. En la misma región, la erosión continua del agua ha dado por resultado la formación de cavernas. Los sumideros se han agrandado, y algunos combinado para formar sumideros de valle. El agua subterránea, filtrándose de los sumideros a la formación calcárea, ha disuelto gran cantidad de caliza. Se formaron cavernas; las que se muestran están conectadas por pasajes horizontales.

En los Estados Unidos existen varias formaciones calizas con numerosas cuevas. Aquí se ve un mapa que muestra los sistemas del enlace Flint y la de la Caverna del Mamut, en Kentucky. La zona baja que lo rodea está surcada por sumideros, y hay un río subterráneo.

trato de yeso (sulfato de calcio hidratado $CaSO_4 \cdot 2H_2O$), otra clase soluble de roca. En algunos lugares, este estrato inclinado de caliza ha sido ahuecado, formando vastas cámaras; la mayor tiene 1.200 metros de largo, 190 de ancho y 100 de alto.

En la región Karst, de Dalmacia, Yugoslavia, los estratos de caliza están plegados. En consecuencia, las cavernas presentan una gran variedad de formas. La superficie del suelo está perforada con numerosos sumideros que, en muchos casos, tienen flancos con paredes escarpadas, a menudo acanaladas o estriadas por el agua que cuela sobre ellas debido a la fuerte precipitación pluvial que hay en la región. Los mayores sumideros o sumideros de valle presentan con frecuencia fondos planos. Los ríos que emergen de sus costados corren sobre estas superficies llanas, para desaparecer de nuevo en una cueva en el otro extremo del valle. En algunos lugares, sumideros alargados o cadenas de sumideros marcan el curso de una corriente de agua subterránea. Se emplea el nombre de *topografía cárstica* para describir paisajes que presentan las mismas características que la región de Karst.

Si los sumideros y sumideros de valle se agrandan más aún, sólo quedarán pequeños restos del área de meseta interpuesta. Con el tiempo, la región mostrará colinas aisladas y esparcidas, cribadas con cavernas; en Puerto Rico se las llama colinas de parva, o de pepino. Formaciones similares aparecen en la región de Cockpit, en Jamaica, y en Cuba al oeste de La Habana. Allí las colinas presentan paredes escarpadas y acanaladas que sugieren una serie de tubos de órgano; a la montaña se la llama Sierra de los Órganos.

Las cavernas calizas pueden ser secas o húmedas; ello depende de las precipitaciones pluviales en la región exterior. Hay cierto retardo, que llega a días o, inclusive, a semanas, entre una lluvia fuera de la cueva y la aparición de agua dentro de ella.

El piso de muchas cavernas de caliza está cubierto con una espesa capa de una arcilla rojiza o amarillenta, infiltrada desde arriba; representa las partículas arcillosas que quedaron cuando la caliza que las contenía se disolvió en la superficie o cerca de ella. Algunos pasadizos de estas grutas están completamente llenos de dichos depósitos, y deben excavarse antes de que puedan penetrar a ellos los visitantes.

Es dudoso que estas cavernas se derrumben alguna vez, si bien hay quienes piensan que de esa manera se forman los puentes naturales. No existe un solo puente natural en la región de la Caverna del Mamut, y pocos en otras zonas calizas. Es probable que paredes y techos de las cavernas, en vez de desmoronarse, se disuelvan gradualmente en el trascurso del tiempo, de modo que, en última instancia, sólo quedarán restos pequeños del techo original. Ocasionalmente, se aflojan y caen bloques rocosos desde éste, pero tales desprendimientos son insignificantes comparados con el volumen total de las cuevas. Los puentes naturales se forman más frecuentemente en la arenisca y otras rocas que en las regiones calizas. En parte son el efecto de la acción erosiva del agua, especialmente por el desgaste que producen los bordes de las corrientes, y, en parte, por la acción de los elementos, incluso los vientos.

No hay dos cavernas calizas iguales; en algunas se entra desde el flanco de una colina; en otras, por el techo, mediante un sumidero. Los pasajes de muchas son bajos; los de otras, lo suficientemente elevados para permitir que una persona alta pase erguida. Algunas no son particularmente notables; otras ofrecen visiones que dejan sin aliento. En la del Mamut, por ejemplo, se llega súbitamente a una barricada al final de un pasaje; desde aquí, se puede ver hacia abajo una inmensa hondonada; por encima, hay una vasta cúpula que trepa casi hasta la superficie de la meseta situada muy arriba. Las paredes, esculpidas en forma de columnas estriadas revelan, hilera sobre hilera, los estratos horizontales de caliza que constituyen toda la región.

DEPÓSITOS DE LAS CAVERNAS

La mayoría de los visitantes de las cuevas calizas se interesa, particularmente, por los depósitos de formas variadas y semejantes a esculturas, asentados con el trascurrir del tiempo; se denominan *espeleotemas*, por dos palabras griegas que significan "asentado en cavernas". Asumen una asombrosa variedad de formas: hay largas y delgadas agujas y destellantes carámbanos suspendidos del techo, gruesos pilares que ascienden desde el suelo, cortinados y colgaduras ondulantes, minaretes y torres con terrazas, cataratas congeladas y una miríada de otras maravillas.

Para explicar la formación de estos depósitos, recordemos que el agua de lluvia se infiltra hacia abajo a través de las rocas y, cargada con bióxido de carbono, reacciona con la caliza y forma bicarbonato de calcio; luego, esta sustancia es arrastrada en solución. El agua saturada de bicarbonato de calcio finalmente desciende hacia el techo de la caverna. A menudo se evapora: libera bióxido de carbono gaseoso y deposita carbonato de calcio, generalmente en forma de calcita.

Ocurre la reacción química siguiente:

$$\underset{\text{bicarbonato de calcio}}{Ca(HCO_3)_2} \quad \underset{\text{da}}{\rightarrow}$$

$$\underset{\text{carbonato de calcio}}{CaCO_3} + \underset{\text{agua}}{H_2O} + \underset{\text{bióxido de carbono}}{CO_2}$$

El depósito de carbonato de calcio toma la forma de un pequeño círculo de piedra. Otras gotas también se depositan allí y, con el tiempo, comienza a descender desde el techo un tubo hueco de piedra que algunos llaman, con bastante propiedad, bombilla. Puede alcanzar más o menos un metro de largo. Con frecuencia, el agua pasa por fuera del tubo hueco y también por el centro. Con el tiempo habrá una formación de calcita en forma de carámbano. Los carámbanos pétreos pueden formar una hilera recta bajo una juntura del techo de la cueva, y alcanzar por lo general varios metros de longitud.

ESTALACTITAS Y ESTALAGMITAS

Todas las formaciones de carbonato de calcio suspendidas del techo de las cavernas se denominan *estalactitas:* del griego *stalaktos*, que significa "manando en gotas". Algunas desarrollan ramas retorcidas que crecen, a veces, hacia arriba… nadie sabe por qué. A estas formaciones ramificadas se las llama *helictitas*, del griego *helix*, que significa "espiral". Allí donde existe un flujo de agua más o menos constante a lo largo de una grieta estrecha del techo, el depósito pétreo puede tomar la forma de una cortina o colgante. A veces, las láminas suspendidas tienen vetas de color castaño causadas por la presencia de limonita, o herrumbre de hierro.

En vez de evaporarse en el techo de la caverna, una gota de agua puede caer al piso justo debajo de un carámbano de estalactita. Al llegar al suelo, la gota salpica y deja su depósito de carbonato de calcio en una zona comparativamente amplia. En este sitio caen otras gotas. Con el tiempo, se genera un depósito romo y redondeado llamado *estalagmita*, del griego *stalagmos*, que significa "un goteo". Tanto las formaciones de estalactitas como las de estalagmitas se llaman *gutolitos*.

En el trascurso de los tiempos, una estalactita y una estalagmita que se ha desarrollado debajo de la primera pueden unirse, formando una *columna*. Como las estalactitas se desarrollan con frecuencia en una hilera que coincide con una grieta en el techo de la caverna, habrá una fila correspondiente de columnas que, con el tiempo, pueden convertirse en una *empalizada* y, finalmente, en una pared que divide la caverna en dos cámaras.

El agua que contiene el bicarbonato de calcio en solución puede abrirse camino hacia abajo, proyectando salientes en las cuevas, evaporándose gradualmente, dejando una formación adecuadamente conocida como *cascada congelada*. Los depósitos de esta clase se denominan *roca de flujo*. En algunos casos, charcos en el piso de la cueva pueden desbordar y originar bordes de calcita: la formación resultante se conoce como *roca calcárea de borde*.

Un poco de materia extraña en un charco puede servir de núcleo para estratos sucesivos de calcita. Con el tiempo se formará una redondeada perla de caverna en la laguna. En casos bastante raros, se puede formar una película de calcita en torno de una burbuja del charco, lo que da una burbuja de calcita que flota en agua; es un dije exquisito, tan frágil que se desmenuza al más ligero toque.

Los depósitos de carbonato de calcio en las cavernas consisten generalmente en calcita. Sin embargo, a veces el carbonato de calcio forma cristales de aragonita a medida que deja la solución. Estos cristales pueden combinarse semejando un hermoso ramillete de flores, conocido como *antodita*, de la palabra griega *anthos*, que significa "flor".

A veces, el yeso puede ser arrastrado a la caverna a medida que rezuma el agua y dejar depósitos en sus superficies cuando aquélla se evapora. Los depósitos de yeso adoptan a veces la forma de rosetones elaborados, de varios centímetros de diámetro y, con frecuencia, de exquisita belleza. También puede encontrarse epsomita, o sulfato de magnesio ($MgSO_4 \cdot 7H_2O$), en forma de acumulaciones de finos cristales fibrosos que revisten las paredes.

OTRAS CLASES DE CAVERNAS

Las grutas también se forman por la acción disolvente del agua en otras formaciones rocosas: dolomita y mármol, estrechamente relacionadas con la piedra caliza y también con el yeso. Dichas grutas son insignificantes en número, en comparación con las calcáreas, y no es preciso tratarlas aquí.

Una cantidad considerable de cavernas se forma por salientes colgantes. En la Figura 3 mostramos la sección trasversal de dos cuevas de este tipo: la superior es bastante pequeña, pero lo suficientemente grande para brindar a una persona protección de la lluvia. Ha resultado de la desintegración de alguna formación débil, tal como pizarra o caliza, o inclusive de arenisca pobremente consolidada, recubierta con un estrato más resistente, como arenisca endurecida. Este estrato es lo suficientemente fuerte para proyectarse hacia el exterior, aun cuando se lo prive de apoyo por debajo.

La caverna inferior que se muestra en el diagrama penetra algo más adentro de la ladera, pero pronto llega a su fin. Puede originarse por la acción del viento, que barre pequeñas partículas de tierra no bien se desprenden de los lados de la caverna, o ser resultado de la infiltración de agua, de madrigueras de animales o del trabajo humano.

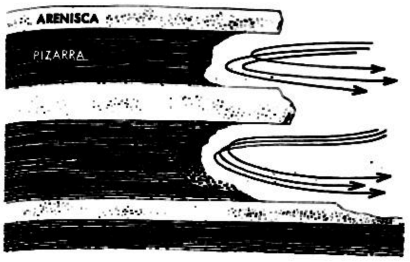


Fig. 3. Dos cavernas formadas por la desintegración de la pizarra, formación rocosa débil entre estratos de arenisca, que es más dura. Las cavernas pueden haber sido horadadas por obra del viento o del agua, o por animales.

Fig. 4. La acción turbulenta de la salpicadura por debajo de una caída de agua puede abrir cuevas, como indica el diagrama. La pizarra, relativamente blanda, ha sido más socavada que la arenisca, más dura. La Caverna de los Vientos, en las Cataratas del Niágara, Nueva York, es un ejemplo de este tipo.

Cierto número de cuevas, tales como la que se muestra en la Figura 4, se ha formado por la acción turbulenta de la salpicadura bajo una caída de agua; la Caverna de los Vientos, en las Cataratas del Niágara, es un ejemplo. Los visitantes pueden descender a ella mediante un ascensor y pararse tras la rugiente cortina líquida. De vez en cuando, enormes bloques rocosos se desprenden de las salientes de arriba, causando ligeros temblores.

Cierto número de cavernas es resultado de la acción de las olas oceánicas; aparecen a lo largo de costas donde las rocas son débiles, están agrietadas, o visiblemente unidas, y por tanto las afecta fácilmente el embate de las olas. Las cavernas marinas formadas de esta suerte usualmente no se extienden mucho tierra adentro. La Gruta Azul de Capri, un islote en la bahía de Nápoles, Italia, es la más famosa de todas las cavernas marinas y una de las más grandes. Tiene 50 metros de profundidad, 30 de ancho y 15 de alto. El techo de la entrada sólo está a un metro por encima del agua. El sol llena la gruta con una deslumbrante luz azul. Muchas otras cavernas aparecen a lo largo de la zona costera de esta región. Todas son formaciones calcáreas erosionadas por la acción de las olas y no desintegradas, como es el caso de la mayoría de las cuevas de caliza.

La Gruta de Fingal, en la isla deshabitada de Staffa, frente a la costa escocesa, es otra caverna marina muy conocida. Labrada en una formación basáltica, sus paredes y boca de acceso parecen estar sostenidas por columnas de basalto. La caverna está a nivel del mar y se penetra en ella por una amplia abertura abovedada de 20 metros de alto y 70 de ancho. Se ha formado a lo largo de una serie de planos de juntura. Llega unos 15 metros adentro y luego termina abruptamente.

Bajo una colada de lava se forma otro tipo de caverna que se produce cuando la corteza exterior se endurece, mientras el interior todavía está fundido. Al fluir este último material queda el techo, formándose una larga cueva, o túnel, en donde cuelgan chorreaduras de lava solidificada. Estas cavernas son comunes en las islas Hawai.

A veces aparecen cavernas en los bancos de arena y arcilla. La tierra herbosa de arriba forma un techo sobre las cavidades abiertas en el suelo flojo subyacente. Con frecuencia, las golondrinas de los acantilados hacen sus nidos bajo las salientes. Las cuevas de esta clase son a menudo bastante grandes y tienen tendencia a desmoronarse súbitamente.

EXPLORADORES DE CAVERNAS

La eterna seducción de las cavernas, cualquiera sea su tipo, atrae una cantidad considerable de osados exploradores –hombres, mujeres, hasta niños– que en algunos países se llaman *spelunkers*, del latín *spelunca*, caverna. Un spelunker verdadero desdeña las rutas de las caver-

nas comerciales, cuidadosamente trazadas para complacer a los visitantes; explora pasajes poco conocidos de esas cuevas y de otras que no están bien exploradas, o han sido relativamente poco cartografiadas.

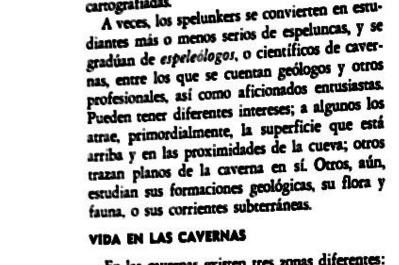
A veces, los spelunkers se convierten en estudiantes más o menos serios de espeluncas, y se gradúan de *espeleólogos*, o científicos de cavernas, entre los que se cuentan geólogos y otros profesionales, así como aficionados entusiastas. Pueden tener diferentes intereses; a algunos los atrae, primordialmente, la superficie que está arriba y en las proximidades de la cueva; otros trazan planos de la caverna en sí. Otros, aún, estudian sus formaciones geológicas, su flora y fauna, o sus corrientes subterráneas.

VIDA EN LAS CAVERNAS

En las cavernas existen tres zonas diferentes: la que está inmediatamente después del acceso; otra de penumbra, en la que penetra cierta cantidad de luz, y otra sumida en oscuridad absoluta. Cada una de estas áreas ofrece un tipo de ambiente particular para los seres que en ella viven.

El pájaro petrolero, o guácharo, anida en las cavernas de Trinidad. Son tan oscuras que las aves deben confiar en su sonar natural para encontrar el camino.

Russ Kinne, Photo Researchers



En las dos zonas exteriores se encuentran cantidades considerables de plantas y animales. Entre las formas vegetales hay algas, líquenes, musgos y helechos. Los animales incluyen osos, comadrejas, mapaches, roedores, murciélagos y aves. Algunos pueden vivir muy adentro, inclusive en la zona de completa oscuridad. Cantidad considerable de animales que habitan en la zona de penumbra nacen y se crían allí, y muchos también mueren en el mismo sitio. Con frecuencia se han encontrado en esas zonas restos de animales extinguidos, así como de formas vivientes.

Quizá los residentes más conocidos de las cavernas sean los pequeños mamíferos voladores llamados murciélagos. Aquí se encuentran varias especies diferentes, pendiendo cabeza abajo del techo o de las partes más elevadas de las paredes. Algunas recurren a la caverna para hibernar; otras la habitan, dejándola a intervalos regulares para buscar alimento.

En muchas cuevas el excremento de murciélago forma depósitos profundos en el piso. Se lo llama *guano*, y es una fuente importante de fertilizante comercial. Durante la Guerra Civil se empleó el guano de murciélago de la Caverna del Mamut para elaborar salitre, un nitrato empleado en la fabricación de explosivos. Todavía se muestra a los visitantes los antiguos tanques y tuberías de madera que se utilizaban en el proceso de manufactura.

Entre los roedores cavernícolas está la rata de caverna, *Neotoma magister*, que construye su nido con corteza de árbol desmenuzada y sustancias similares, muy dentro de la cueva, a veces en el área de oscuridad total. *Neotoma magister* abandona su morada para ir a buscar las semillas y nueces con que se alimenta.

De los pájaros que hacen aquí su nido quizás el más notable sea el guácharo o pájaro petrolero, de Venezuela. Estas aves, con una envergadura de alas de más de un metro, viven en las zonas oscuras, muy dentro de las cuevas. Salen durante el crepúsculo a buscar alimento y vuelven antes del amanecer. A veces, en una sola caverna habitan varios miles de guácharos, los que producen un ruido espantoso cuando salen. Otros pájaros que moran en cavernas son búhos, papamoscas y grajos.

Buena cantidad de material nutritivo entra desde el exterior a la zona de completa oscuridad. Los cursos de agua llevan materias orgánicas provenientes de las plantas. Cuando el viento circula por la caverna acarrea consigo esporas de hongos. Pueden germinar semillas llevadas al interior por animales; después de unos pocos días de crecimiento enfermizo los almácigos mueren y se agregan al depósito de alimento orgánico. Todos estos materiales constituyen la base de un ciclo alimentario que implica una variada población de formas animales, y unas pocas vegetales, las que trascurren toda su vida bajo tierra.

En términos generales, en la profundidad de las cavernas sólo se encuentran plantas que pueden subsistir en ausencia de luz solar (tales como hongos y bacterias). Los hongos prosperan en esas condiciones; en verdad, en muchas zonas se los cultiva en cavernas. También abundan mohos de una especie de *Penicillium*, los que se emplean para hacer madurar el queso roquefort en cavernas del sur de Francia. Las plantas que requieren luz solar a veces germinan en la zona totalmente oscura, pero no pueden alcanzar la madurez ni reproducirse.

Entre los animales que trascurren su vida en el área de total oscuridad hay gusanos, peces, anfibios, insectos, arañas y crustáceos. En general, son depredadores que comen a otros animales o coprófagos –que se alimentan de materia orgánica en descomposición.

Un interesante animal acuático que se encuentra muy adentro en las cuevas es el pez ciego (*Amblyopsis spelaeus*), de las cavernas de Kentucky. El *Amblyopsis* se ha vuelto por completo etiolado, o blanqueado, y sus ojos se redujeron a meros vestigios sin función. Para compensar la carencia de visión, está provisto de papilas, o protuberancias, sumamente sensibles, dispuestas como arrugas en el frente y costados de la cabeza. En las cuevas hay también langostinos ciegos.

Los anfibios comprenden varias especies de salamandras y ranas. Ciertas salamandras cavernícolas presentan ojos degenerados, casi o enteramente cubiertos por párpados fusionados. Otras, que se hallan en varias cavernas de Australia, carecen de ellos. Algunas salamandras son blancas; otras, moteadas.

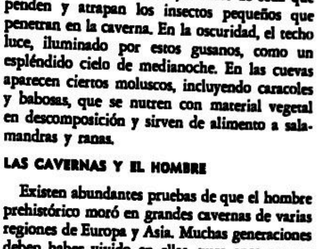
Los principales habitantes no acuáticos de las zonas completamente oscuras son los artrópodos, incluyendo insectos, arañas, crustáceos y su linaje; también ciempiés, gorgojos, polillas, escarabajos, moscas, grillos y pulgas. Entre los insectos más interesantes está la mosca llamada *Bolitophila luminosa*: sus larvas, conocidas como gusanos refulgentes, viven en las cavernas de Waitomo, Nueva Zelanda, en colonias densas que cuelgan del techo. Tejen hilos de seda que penden y atrapan los insectos pequeños que penetran en la caverna. En la oscuridad, el techo luce, iluminado por estos gusanos, como un espléndido cielo de medianoche. En las cuevas aparecen ciertos moluscos, incluyendo caracoles y babosas, que se nutren con material vegetal en descomposición y sirven de alimento a salamandras y ranas.

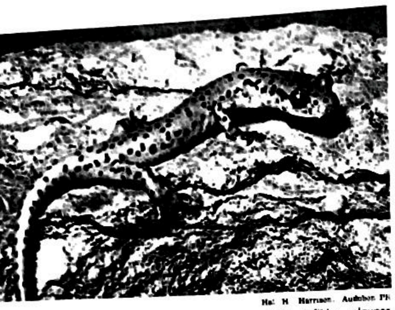
LAS CAVERNAS Y EL HOMBRE

Existen abundantes pruebas de que el hombre prehistórico moró en grandes cavernas de varias regiones de Europa y Asia. Muchas generaciones deben haber vivido en ellas, pues encontramos sus herramientas, utensilios y desechos caseros apilados capa sobre capa, juntamente con los huesos de los animales que cazaron. Las paredes de algunas de las cavernas que habitó el hombre prehistórico contienen bellas muestras de arte, siendo particularmente notables las pinturas ru-

La vista es inútil para muchos animales cavernícolas. Este pez ciego, que habita en las corrientes de agua subterránea, sólo muestra vestigios de órbitas oculares.

Russ Kinne Pr





Hal H. Harrison, Audubon Ph

También viven en las cavernas ciertos anfibios, algunos hasta en las oquedades más profundas. He aquí una salamandra cavernícola moteada. Otras especies de salamandra que habitan cuevas son típicamente blancas.

pestres de Altamira, España, y de Font-de-Gaume, Francia.

Los antiguos griegos escribieron sobre ciertos pueblos primitivos, a los que llamaron *trogloditas*, o habitantes de las cavernas. Vivían en las de muchos lugares, en torno de las costas del Mediterráneo, especialmente en Libia y, hacia el este, hasta Egipto. En algunos casos, pueblos antiguos construyeron ciudades subterráneas, talladas en la roca sólida: entre éstas se cuentan Edrei, en Siria, y Petra, en Jordania. Los antiguos chinos también utilizaban cavernas como morada.

Hoy en día, aún hay gente que las habita; en ciertas partes de China y Austria los habitantes excavan cámaras en depósitos de loess. También hay cavernícolas en diversas partes de Francia, España e Italia.

Quizá las cavernas sirvieron como refugio durante millones de años. Cuando los romanos atacaron las Galias, los habitantes a menudo se escondían en aquéllas; siglos después fueron asilos al producirse la invasión de los sarracenos. Los anales medievales abundan en relatos de refugios de este tipo. En los Estados Unidos se dice que la Caverna del Mamut se usó como "estación" del llamado ferrocarril subterráneo, que ayudaba a los esclavos a escapar de los Estados del Sur, antes de la Guerra de Secesión.

También se las ha utilizado ampliamente como lugares de culto religioso. En Egipto, a veces se esculpían templos en la roca sólida y se los dotaba con accesos ornamentados. En la India existen maravillosos templos en cavernas, que se remontan a tiempos antiguos, particularmente en Elefanta, isla de Bombay. En cavernas chinas se establecieron muchos templos budistas. En Europa occidental, varios monasterios famosos fueron originalmente ermitas subterráneas, ulteriormente agrandadas.

Jim Hubbard, Photo Researchers

Los moradores de las cavernas se conocen como *trogloditas*. Esta foto muestra un grupo de gitanos españoles trogloditas cerca de sus viviendas.

Se les ha encontrado varios usos más y se las empleó como espaciosas prisiones. Los cerveceros de San Luis, Misurí, en un tiempo almacenaron su producto en cavernas. Por lo menos, una sirve como túnel de ferrocarril: el Túnel Natural, utilizado por el Southern Railroad en el sudoeste de Virginia. Ocasionalmente, las corrientes subterráneas que fluyen en las cavernas han provisto de agua a comunidades. Empero, esta fuente tiene que emplearse con cautela, puesto que los sumideros sirven como fosas para basura en algunas zonas.

SE TERMINO LA IMPRESION DE ESTA OBRA EN MAYO DE 1983 EN IMPRESORA Y EDITORA MEXICANA, S.A. DE C.V., EN SAN MATEO TECOLOAPAN EDO. DE MEXICO
SE IMPRIMIERON 40,000 EJEMPLARES.